

(19) 日本国特許庁 (J P)

(12) 特 許 公 報 (B 2)

(11) 特許番号

第2927350号

(45) 発行日 平成11年(1999) 7月28日

(24) 登録日 平成11年(1999) 5月14日

| (51) Int.Cl. <sup>8</sup> | 識別記号 | F I     |         |
|---------------------------|------|---------|---------|
| G 0 6 T                   | 5/20 | G 0 6 F | 15/68   |
|                           | 7/00 |         | 15/62   |
|                           |      |         | 15/70   |
|                           |      |         | 4 0 0 A |
|                           |      |         | 4 1 5   |
|                           |      |         | 4 6 0 Z |

請求項の数31(全 25 頁)

|           |                   |  |   |
|-----------|-------------------|--|---|
| (21) 出願番号 | 特願平9-95318        | (73) 特許権者                                | 396001980<br>株式会社モノリス<br>東京都港区麻布十番1丁目7番3号     |
| (22) 出願日  | 平成9年(1997) 3月27日  | (72) 発明者                                 | 品川 嘉久<br>東京都江戸川区西葛西5-10-26-204                |
| (65) 公開番号 | 特開平10-269355      | (72) 発明者                                 | 國井 利春<br>東京都文京区本郷1-25-21 ドムス本郷602             |
| (43) 公開日  | 平成10年(1998) 10月9日 | (74) 代理人                                 | 弁理士 関 正治 (外1名)                                |
| 審査請求日     | 平成10年(1998) 3月17日 | 審査官                                      | 後藤 彰  |
|           |                   | (56) 参考文献                                | 特開 平9-97334 (J P, A)<br>特開 昭62-80768 (J P, A) |
|           |                   | (58) 調査した分野(Int.Cl. <sup>8</sup> , D B名) | G06T 5/00 - 7/60<br>G06T 15/00 - 15/40        |

(54) 【発明の名称】 多重解像度フィルタ処理方法およびその方法を利用することのできる画像マッチング方法

1

(57) 【特許請求の範囲】

【請求項1】 第一の画像に対し、二次元的な探索を行って特異点を検出する検出工程と、  
検出された特異点を抽出して第一の画像よりも解像度の低い第二の画像を生成する生成工程と、  
を含むことを特徴とする多重解像度フィルタ処理方法。

【請求項2】 請求項1に記載の方法において、  
検出工程は、第一の画像を構成する複数のブロックのそれぞれの内部で特異点を探索する多重解像度フィルタ処理方法。

【請求項3】 請求項2に記載の方法において、  
検出工程は、各ブロックのふたつの方向について最大画素値または最小画素値をもつ点を探索することによって特異点を検出する多重解像度フィルタ処理方法。

【請求項4】 請求項3に記載の方法において、

2

検出工程は、前記ふたつの方向の両方について最大画素値となる画素を極大点として検出する多重解像度フィルタ処理方法。

【請求項5】 請求項3に記載の方法において、  
検出工程は、前記ふたつの方向の両方について最小画素値となる画素を極小点として検出する多重解像度フィルタ処理方法。

【請求項6】 請求項3に記載の方法において、  
検出工程は、前記ふたつの方向の一方について最大画素値となるとともに、他方について最小画素値となる画素を鞍点として検出する多重解像度フィルタ処理方法。

【請求項7】 請求項3に記載の方法において、前記ブロックはそれぞれ縦横2個ずつの合計4画素から構成され、これら4画素が極大点、極小点、および二種類の鞍点のいずれかに分類される多重解像度フィルタ処理方

10

法。

【請求項8】 請求項2～7のいずれかに記載の方法において、

生成工程は、各ブロックの内部で検出された特異点の画像でそのブロックの画像を代表させることにより、画像の解像度を落とす多重解像度フィルタ処理方法。

【請求項9】 請求項2～8のいずれかに記載の方法において、

生成工程は、各ブロックの内部で検出された特異点の種類ごとに第二の画像を生成する多重解像度フィルタ処理方法。

【請求項10】 始点画像に多重解像度特異点フィルタを施して解像度の異なる一連の始点階層画像を生成する第一工程と、

終点画像に多重解像度特異点フィルタを施して解像度の異なる一連の終点階層画像を生成する第二工程と、

始点階層画像と終点階層画像のマッチングを解像度レベルの階層の中で計算する第三工程と、

を含むことを特徴とする画像マッチング方法。

【請求項11】 請求項10に記載の方法において、第三工程は、始点階層画像中のある解像度レベルの画像と、終点階層画像中の同一解像度レベルの画像間の写像を、他の所定の解像度レベルにおける写像を考慮しながら決定する画像マッチング方法。

【請求項12】 請求項11に記載の方法において、第三工程は、前記写像を決定する際、前記他の所定の解像度レベルにおける写像を拘束条件として用いる画像マッチング方法。

【請求項13】 請求項11、12のいずれかに記載の方法において、

前記所定の解像度レベルは、現在写像を決定しようとしている解像度レベルよりも粗い解像度レベルである画像マッチング方法。

【請求項14】 請求項13に記載の方法において、前記所定の解像度レベルは、現在写像を決定しようとしている解像度レベルよりも1レベルだけ粗い解像度レベルである画像マッチング方法。

【請求項15】 請求項11～14のいずれかに記載の方法において、

第三工程は、まず最も粗い解像度レベルにおいて写像を決定し、以降、次第に解像度レベルが細くなる順に写像を決定していく画像マッチング方法。

【請求項16】 請求項11～15のいずれかに記載の方法において、

第三工程は、全単射条件を満たすよう前記写像を決定する画像マッチング方法。

【請求項17】 請求項16に記載の方法において、第三工程は、全単射条件に緩和条件を加味する画像マッチング方法。

【請求項18】 請求項17に記載の方法において、

前記緩和条件は、写像が収縮写像になることを認める旨の条件である画像マッチング方法。

【請求項19】 請求項11～18のいずれかに記載の方法において、

第一工程および第二工程は、特異点の種類ごとにそれぞれ一連の始点階層画像および終点階層画像を生成し、第三工程は特異点の種類ごとに写像を求める画像マッチング方法。

【請求項20】 請求項19に記載の方法において、第三工程は、ある特異点に関して写像を求めるとき、同一の解像度レベルにおいてすでに求められている他の種類の特異点に関する写像を考慮する画像マッチング方法。

【請求項21】 請求項20に記載の方法において、前記他の種類の特異点に関する写像との類似度が高くなるよう条件を課して写像を求める画像マッチング方法。

【請求項22】 請求項10～21のいずれかに記載の方法において、

第三工程は、複数のマッチング評価項目のそれぞれについて評価式を設け、それらの評価式を統合して総合評価式を定義し、その総合評価式の極値付近に注目して最適マッチングを探索することを特徴とする画像マッチング方法。

【請求項23】 請求項22に記載の方法において、前記総合評価式は、評価式の少なくともひとつに係数パラメータを掛けたうえでそれらの評価式の総和として定義される画像マッチング方法。

【請求項24】 請求項23に記載の方法において、前記評価式はそれぞれ評価結果が良好なほど値が小さくなり、前記総合評価式がとりうる極小値が最小になるよう前記パラメータが自動決定される画像マッチング方法。

【請求項25】 請求項23に記載の方法において、前記評価式はそれぞれ評価結果が良好なほど値が大きくなり、前記総合評価式がとりうる極大値が最大になるよう前記パラメータが自動決定される画像マッチング方法。

【請求項26】 請求項23に記載の方法において、いずれかの評価式がほぼ極値をとる状態を検出し、前記パラメータを自動決定する画像マッチング方法。

【請求項27】 請求項22に記載の方法において、画素値に関連する第一評価式と画素の位置に関連する第二評価式の線形和で総合評価式を定義し、

少なくとも第一評価式の係数パラメータを変化させながら総合評価式がほぼ極値をとるときの第一評価式の値を記録し、

第一評価式がほぼ極値をとるときの係数パラメータを検出し、

その係数パラメータを以降の評価に固定的に用いる画像マッチング方法。

【請求項28】 始点画像と終点画像のマッチングのために、複数のマッチング評価項目のそれぞれについて評価式を設け、それらの評価式を統合して総合評価式を定義し、その総合評価式の極値付近に注目して最適マッチングを探索することを特徴とする画像マッチング方法であって、

前記総合評価式は、評価式の少なくともひとつに係数パラメータを掛けただけでそれらの評価式の総和として定義され、

前記評価式はそれぞれ評価結果が良好なほど値が小さくなり、前記総合評価式のとり極小値が最小になるよう前記パラメータが自動決定される画像マッチング方法。

【請求項29】 始点画像と終点画像のマッチングのために、複数のマッチング評価項目のそれぞれについて評価式を設け、それらの評価式を統合して総合評価式を定義し、その総合評価式の極値付近に注目して最適マッチングを探索することを特徴とする画像マッチング方法であって、

前記総合評価式は、評価式の少なくともひとつに係数パラメータを掛けただけでそれらの評価式の総和として定義され、

前記評価式はそれぞれ評価結果が良好なほど値が大きくなり、前記総合評価式のとり極大値が最大になるよう前記パラメータが自動決定される画像マッチング方法。

【請求項30】 始点画像と終点画像のマッチングのために、複数のマッチング評価項目のそれぞれについて評価式を設け、それらの評価式を統合して総合評価式を定義し、その総合評価式の極値付近に注目して最適マッチングを探索することを特徴とする画像マッチング方法であって、

前記総合評価式は、評価式の少なくともひとつに係数パラメータを掛けただけでそれらの評価式の総和として定義され、

いずれかの評価式がほぼ極値をとる状態を検出し、前記パラメータを自動決定する画像マッチング方法。

【請求項31】 始点画像と終点画像のマッチングのために、複数のマッチング評価項目のそれぞれについて評価式を設け、それらの評価式を統合して総合評価式を定義し、その総合評価式の極値付近に注目して最適マッチングを探索することを特徴とする画像マッチング方法において、

画素値に関連する第一評価式と画素の位置に関連する第二評価式の線形和で総合評価式を定義し、

少なくとも第一評価式の係数パラメータを変化させながら総合評価式がほぼ極値をとるときの第一評価式の値を記録し、

第一評価式がほぼ極値をとるときの係数パラメータを検出し、

その係数パラメータを以降の評価に固定的に用いる画像マッチング方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】この発明は、多重解像度フィルタ処理方法および画像マッチング方法に関する。この発明は特に、多重解像度フィルタによって新たな画像を生成する方法、およびその方法を利用することの可能な画像マッチング方法に関する。

【0002】

【従来の技術】ふたつの画像の自動的なマッチング、つまり画像領域や画素どうしの対応付けは、コンピュータビジョンやコンピュータグラフィックスにおける最も難しくかつ重要なテーマのひとつである。例えば、あるオブジェクトに関して異なる視点からの画像間でマッチングがとれれば、他の視点からの画像を生成することができる。右目画像と左目画像のマッチングが計算できれば、立体画像を用いた写真測量も可能である。顔の画像のモデルと他の顔の画像のマッチングがとれたとき、目、鼻、口といった特徴的な顔の部分の抽出することができる。例えば人の顔と猫の顔の画像間でマッチングが正確にとれたとき、それらの中割画像を自動的に生成することでモーフィングを完全自動化することができる。

【0003】しかし従来一般に、ふたつの画像間の対応点は人がいちいち指定しなければならず、多大な作業工数を要した。この問題を解消するために数多くの対応点自動検出方法が提案されている。例えば、エッジ検出線を用いることによって対応点の候補の数を減らす考えがある。しかし、その場合でも処理はきわめて複雑である。複雑さを低減するために、左目画像の各点の座標は通常右目画像でもほぼ同じ位置にあると想定される。しかし、こうした制約を設けると、大域的特徴及び局所の特徴を同時に満たすマッチングをとることは非常に困難になる。

【0004】ポリウムレンダリングでは、ボクセルを構成するために一連の断面画像が用いられる。この場合、従来一般に、上方の断面画像における画素が下方の断面画像の同一箇所にある画素と対応すると仮定され、これらの画素のペアが内挿計算に用いられる。このようにきわめて単純な方法を用いるため、連続する断面間の距離が遠く、オブジェクトの断面形状が大きく変化する場合、ポリウムレンダリングで構築されたオブジェクトは不明瞭になりがちである。

【0005】立体写真測量法など、エッジの検出を利用するマッチングアルゴリズムも多い。しかしこの場合、結果的に得られる対応点の数が少ないため、マッチングのとれた対応点間のギャップを埋めるべく、ディスパリティの値を内挿計算しなければならない。一般にあらゆるエッジ検出器は、それらが用いる局所的なウィンドウの中で画素の輝度に変化したとき、これが本当にエッジの存在を示唆するかどうかを判断することが難しい。エ

ッジ検出器は、本来的にすべてハイパスフィルタであり、エッジと同時にノイズも拾ってしまう。

【0006】さらに別の手法として、オブティカルフローが知られている。二枚の画像が与えられたとき、オブティカルフローでは画像内のオブジェクト（剛体）の動きを検出する。その際、オブジェクトの各画素の輝度は変化しないと仮定する。オブティカルフローでは例えば（ $u$ 、 $v$ ）のベクトル場の滑らかさといった、いくつかの付加的な条件とともに、各画素の動きベクトル（ $u$ 、 $v$ ）を計算する。しかし、オブティカルフローでは画像間の大域的な対応関係を検出することはできない。画素の輝度の局所的な変化に注目するのみであり、画像の変位が大きい場合、システムの誤差は顕著になる。

【0007】画像の大域的な相違を認識するために、多重解像度フィルタも数多く提案されてきた。それらは線形フィルタと非線形フィルタに分類される。前者の例としてウェーブレットがあるが、線形フィルタは一般に、画像マッチングにはさして有用ではない。なぜなら、極値をとる画素の輝度に関する情報がそれらの位置情報とともに次第に不鮮明になるためである。図1(a)と図1(b)はそれぞれ図19(a)と図19(b)に示す顔の画像に対して平均化フィルタを適用した結果を示している。同図のごとく、極値をとる画素の輝度が平均化によって次第に薄れるとともに、位置も平均化の影響でシフトしていく。その結果、目（輝度の極小点）の輝度や位置の情報は、このような粗い解像度レベルで曖昧になり、この解像度では正しいマッチングを計算することができない。したがって、粗い解像度レベルを設けるのが大域的なマッチングのためでありながら、ここで得られたマッチングは画像の本当の特徴（目、つまり極小点）に正確に対応しない。より精細な解像度レベルで目が鮮明に現れたとしても、大域的なマッチングをとる際に混入した誤差は、もはや取り返しがつかない。入力画像にスムージング処理を加えることにより、テクスチャ領域のステレオ情報が落ちてしまうこともすでに指摘されている。

【0008】一方、最近地形学の分野で利用されはじめた非線形フィルタとして一次元の「ふるい（sieve）」演算子がある。この演算子は、所定の大きさの一次元ウィンドウ内の極小値（または極大値）を選択することにより、縮尺と空間の因果関係を保存しながら画像にスムージング処理を加える。その結果得られる画像は元の画像と同じ大きさであるが、小さな波の成分が取り除かれているため、より単純になる。画像の情報を落とすという点で、この演算子は広い意味での「多重解像度フィルタ」に分類することはできるが、実際にはウェーブレットのように画像の解像度を変えながら画像を階層化するわけではなく（つまり狭い意味での多重解像度フィルタではなく）、画像間の対応の検出には利用できない。

【0009】

【発明が解決しようとする課題】以上をまとめれば以下の課題が認められる。

【0010】1. 画像の特徴を正確に、かつ比較的簡単な処理で把握する画像処理方法が乏しかった。特に、特徴のある点に関する情報、例えば画素値や位置を維持しながら特徴を抽出できる画像処理方法に関する有効な提案が少なかった。

2. 画像の特徴をもとに対応点を自動検出する場合、一般に処理が複雑であるか、ノイズ耐性が低いなどの欠点があった。また、処理に際していろいろな制約を設ける必要があり、大域的特徴及び局所的特徴を同時に満たすマッチングをとることが困難だった。

3. 画像の大域的な相違または特徴を認識するために多重解像度フィルタを導入しても、そのフィルタが線形フィルタの場合、画素の輝度情報と位置情報が曖昧になった。その結果、対応点の把握が不正確になりやすかった。非線形フィルタである一次元ふるい演算子は画像を階層化しないため、画像間の対応点の検出には利用できなかった。

4. これらの結果、対応点を正しく把握しようとするれば、結局人手による指定に頼るほか有効な手だてがなかった。

【0011】本発明はこれらの課題の解決を目的としてなされたものであり、画像処理の分野において、画像の特徴の的確な把握を可能にする技術を提供するものである。

【0012】

【課題を解決するための手段】この目的のために本発明のある態様は、新たな多重解像度の画像フィルタを提案する。この多重解像度フィルタは画像から特異点を抽出する。したがって、特異点フィルタともよばれる。特異点とは画像上特徴をもつ点をいう。例として、ある領域において画素値（画素値とは、色番号、輝度値など画像または画素に関する任意の数値を指すものとする）が最大になる極大点、最小になる極小点、ある方向については最大だが別の方向については最小になるような鞍点がある。特異点は位相幾何学上の概念であってもよい。ただし、その他どのような特徴を有してもよい。いかなる性質の点を特異点と考えるかは、本発明にとって本質問題ではない。

【0013】この態様では、多重解像度フィルタを用いた画像処理が行われる。まず検出工程において、第一の画像に対し、二次元的な探索を行って特異点が検出される。つぎに生成工程において、検出された特異点を抽出して第一の画像よりも解像度の低い第二の画像が生成される。第二の画像には第一の画像のもつ特異点が引き継がれる。第二の画像は第一の画像よりも解像度が低いいため、画像の大域的な特徴の把握に好適である。

【0014】本発明の別の態様は特異点フィルタを用いた画像マッチング方法に関する。この態様では、始点画

像と終点画像間のマッチングがとられる。始点画像および終点画像とは、ふたつの画像の区別のために便宜的に与えた名称であり、本質的な違いはない。

【0015】この態様では、第一工程にて、始点画像に特異点フィルタを施して解像度の異なる一連の始点階層画像が生成される。第二工程では、終点画像に特異点フィルタを施して解像度の異なる一連の終点階層画像が生成される。始点階層画像、終点階層画像とは、それぞれ始点画像、終点画像を階層化して得られる画像群をいい、それぞれ最低2枚の画像からなる。つぎに第三工程において、始点階層画像と終点階層画像のマッチングが

解像度レベルの階層の中で計算される。この態様によれば、多重解像度フィルタによって特異点に関連する画像の特徴が抽出され、および/または明確化されるため、マッチングが容易になる。マッチングのための拘束条件は特に必要としない。

【0016】本発明のさらに別の態様も始点画像と終点画像のマッチングに関する。この態様では、予め複数のマッチング評価項目のそれぞれに関して評価式を設け、それらの評価式を統合して総合評価式を定義し、その総合評価式の極値付近に注目して最適マッチングを探索する。総合評価式は、評価式の少なくともひとつに係数パラメータを掛けたりしてそれらの評価式の総和として定義してもよく、その場合、総合評価式またはいずれかの評価式がほぼ極値をとる状態を検出して前記パラメータを決定してもよい。「極値付近」または「ほぼ極値をとる」としたのは、多少誤差を含んでいてもよいためである。多少の誤差は本発明にはさして問題とならない。

【0017】極値自体も前記パラメータに依存するため、極値の挙動、つまり極値の変化の様子をもとに、最適と考えられるパラメータを決定する余地が生じる。この態様はその事実を利用している。この態様によれば、元来調整の困難なパラメータの決定を自動化する途が拓かれる。

【0018】

【発明の実施の形態】最初に[1]で実施の形態の要素技術の詳述し、[2]で処理手順を具体的に説明する。

さらに[3]で実験の結果を報告する。

\*

$$\begin{aligned}
 p_{(i,j)}^{(m,0)} &= \min(\min(p_{(2i,2j)}^{(m+1,0)}, p_{(2i,2j+1)}^{(m+1,0)}), \min(p_{(2i+1,2j)}^{(m+1,0)}, p_{(2i+1,2j+1)}^{(m+1,0)})) \\
 p_{(i,j)}^{(m,1)} &= \max(\min(p_{(2i,2j)}^{(m+1,1)}, p_{(2i,2j+1)}^{(m+1,1)}), \min(p_{(2i+1,2j)}^{(m+1,1)}, p_{(2i+1,2j+1)}^{(m+1,1)})) \\
 p_{(i,j)}^{(m,2)} &= \min(\max(p_{(2i,2j)}^{(m+1,2)}, p_{(2i,2j+1)}^{(m+1,2)}), \max(p_{(2i+1,2j)}^{(m+1,2)}, p_{(2i+1,2j+1)}^{(m+1,2)})) \\
 p_{(i,j)}^{(m,3)} &= \max(\max(p_{(2i,2j)}^{(m+1,3)}, p_{(2i,2j+1)}^{(m+1,3)}), \max(p_{(2i+1,2j)}^{(m+1,3)}, p_{(2i+1,2j+1)}^{(m+1,3)}))
 \end{aligned}$$

\*【0019】[1]要素技術の詳細

[1.1]イントロダクション

特異点フィルタと呼ばれる新たな多重解像度フィルタを導入し、画像間のマッチングを正確に計算する。オブジェクトに関する予備知識は一切不要である。画像間のマッチングの計算は、解像度の階層を進む間、各解像度において計算される。その際、粗いレベルから細かいレベルへと順に解像度の階層を述っていく。計算に必要なパラメータは、人間の視覚システムに似た動的計算によって完全に自動設定される。画像間の対応点を人手で特定する必要はない。

【0020】本実施の形態は、例えば完全に自動的なモーフィング、物体認識、立体写真測距、ポリウムレンダリング、少ないフレームからの滑らかな動画の生成などに応用できる。モーフィングに用いる場合、与えられた画像を自動的に変形することができる。ポリウムレンダリングに用いる場合、断面間の中間的な画像を正確に再構築することができる。断面間の距離が遠く、断面の形状が大きく変化する場合でも同様である。

【0021】[1.2]特異点フィルタの階層

実施の形態に係る多重解像度特異点フィルタは、画像の解像度を落としながら、しかも画像に含まれる各特異点の傾度及び位置を保存することができる。ここで画像の幅をN、高さをMとする。以下簡単のため、 $N=M=2^n$  ( $n$ は自然数)と仮定する。また、区間 $[0, N] \subset \mathbb{R}$ をIと記述する。 $(i, j)$ における画像の画素を $p_{(i,j)}$ と記述する( $i, j \in I$ )。

【0022】ここで多重解像度の階層を導入する。階層化された画像群は多重解像度フィルタで生成される。多重解像度フィルタは、もとの画像に対して二次元的な探索を行って特異点を検出し、検出された特異点を抽出してもとの画像よりも解像度の低い別の画像を生成する。ここで第mレベルにおける各画像のサイズは $2^m \times 2^m$  ( $0 \leq m \leq n$ )とする。特異点フィルタは次の4種類の新たな階層画像をnから下がる方向で再帰的に構築する。

【0023】

【数1】

(式1)

ただしここで、

$$p_{(i,j)}^{(n,0)} = p_{(i,j)}^{(n,1)} = p_{(i,j)}^{(n,2)} = p_{(i,j)}^{(n,3)} = p_{(i,j)} \quad (\text{式2})$$

とする。以降これら4つの画像を副画像（サブイメージ）と呼ぶ。 $\min x \leq t \leq x+1$ ,  $\max x \leq t \leq x+1$  をそれぞれ $\alpha$ 及び $\beta$ と記述すると、副画像はそれぞれ以下のように記述できる。

$$\begin{aligned} [0024] \quad & p^{(\alpha,0)} = \alpha(x) \alpha(y) p^{(\alpha,1,0)} \\ & p^{(\alpha,1)} = \alpha(x) \beta(y) p^{(\alpha,1,1)} \\ & p^{(\alpha,2)} = \beta(x) \alpha(y) p^{(\alpha,1,2)} \\ & p^{(\alpha,3)} = \beta(x) \beta(y) p^{(\alpha,1,3)} \end{aligned}$$

すなわち、これらは $\alpha$ と $\beta$ のテンソル積のようなものと考えられる。副画像はそれぞれ特異点に対応している。これらの式から明かなように、特異点フィルタはもとの画像について $2 \times 2$ 画素で構成されるブロックごとに特異点を検出する。その際、各ブロックのふたつの方向、つまり縦と横について、最大画素値または最小画素値をもつ点を探索する。画素値として、実施の形態では輝度を採用するが、画像に関するいろいろな数値を採用

【0025】特異点フィルタは、各ブロックの内部で検出された特異点の画像（ここでは1画素）でそのブロックの画像（ここでは4画素）を代表させることにより、画像の解像度を落とす。特異点の理論的な傾点からすれば、 $\alpha(x) \alpha(y)$ は極小点を保存し、 $\beta(x) \beta(y)$ は極大点を保存し、 $\alpha(x) \beta(y)$ 及び $\beta(x) \alpha(y)$ は鞍点を保存する。

【0026】はじめに、マッチングをとるべき始点（ソース）画像と終点（デスティネーション）画像に対して別々に特異点フィルタ処理を施し、それぞれ一連の画像群、すなわち始点階層画像と終点階層画像を生成しておく。始点階層画像と終点階層画像は、特異点の種類に対応してそれぞれ4種類ずつ生成される。

【0027】この後、一連の解像度レベルの中で始点階層画像と終点階層画像のマッチングがとられていく。まず $p^{(\alpha,0)}$ を用いて極小点のマッチングがとられる。次に、その結果に基づき、 $p^{(\alpha,1)}$ を用いて鞍点のマッチ

$$f(m,i) : p^{(m,\sigma(i))} \rightarrow q^{(m,\sigma(i))}$$

ここで $\sigma(i) \in \{0, 1, 2, 3\}$ である。

【0032】[1. 3. 1] 全単射

始点画像と終点画像の間のマッチングを写像で表現する場合、その写像は両画像間で全単射条件を満たすべきである。両画像に概念上の優劣はなく、互いの画素が全射

※ ※ 【数2】

※ングがとられ、 $p^{(\alpha,2)}$ を用いて他の鞍点のマッチングがとられる。そして最後に $p^{(\alpha,3)}$ を用いて極大点のマッチングがとられる。

【0028】図1(c)と図1(d)はそれぞれ図1(a)と図1(b)の副画像 $p^{(\alpha,0)}$ を示している。同様に、図1(e)と図1(f)は $p^{(\alpha,1)}$ 、図1(g)と図1(h)は $p^{(\alpha,2)}$ 、図1(i)と図1(j)は $p^{(\alpha,3)}$ をそれぞれ示している。これらの図からわかるとおり、副画像によれば画像の特徴部分のマッチングが容易になる。まず $p^{(\alpha,0)}$ によって目が明確になる。目は顔の中で輝度の極小点だからである。 $p^{(\alpha,1)}$ によれば口が明確になる。口は横方向で輝度が低いからである。 $p^{(\alpha,2)}$ によれば首の両側の縦線が明確になる。最後に、 $p^{(\alpha,3)}$ によって耳や頬の最も明るい点が明確になる。これらは輝度の極大点だからである。

【0029】特異点フィルタによれば画像の特徴が抽出できるため、例えばカメラで撮影された画像の特徴と、予め記録しておきたいいくつかのオブジェクトの特徴を比較することにより、カメラに映った被写体を識別することができる。

【0030】[1. 3] 画像間の写像の計算  
始点画像の位置 $(i, j)$ の画素を $p^{(n)}_{(i,j)}$ と書き、同じく終点画像の位置 $(k, l)$ の画素を $q^{(n)}_{(k,l)}$ で記述する。 $i, j, k, l \in I$ とする。画像間の写像のエネルギー（後述）を定義する。このエネルギーは、始点画像の画素の輝度と終点画像の対応する画素の輝度の差、及び写像の滑らかさによって決まる。最初に最小のエネルギーを持つ $p^{(\alpha,0)}$ と $q^{(\alpha,0)}$ 間の写像 $f^{(\alpha,0)}$ ： $p^{(\alpha,0)} \rightarrow q^{(\alpha,0)}$ が計算される。 $f^{(\alpha,0)}$ に基づき、最小エネルギーを持つ $p^{(\alpha,1)}$ 、 $q^{(\alpha,1)}$ 間の写像 $f^{(\alpha,1)}$ が計算される。この手続は、 $p^{(\alpha,2)}$ と $q^{(\alpha,2)}$ の間の写像 $f^{(\alpha,2)}$ の計算が終了するまで続く。各写像 $f^{(\alpha,i)}$  ( $i = 0, 1, 2, \dots$ )を副写像と呼ぶことにする。 $f^{(\alpha,i)}$ の計算の都合のために、 $i$ の順序は次式のように並べ替えることができる。並べ替えが必要な理由は後述する。

【0031】

【数3】

(式3)

かつ単射で接続されるべきだからである。しかしながら通常の場合とは異なり、ここで構築すべき写像は全単射のデジタル版である。実施の形態では、画素は格子点によって特定される。

【0033】始点副画像（始点画像について設けられた

副画像)から終点副画像(終点画像について設けられた副画像)への写像は、 $f^{(m,s)}: 1/2^{m-s} \times 1/2^{m-s} \rightarrow 1/2^{m-s} \times 1/2^{m-s}$  ( $s=0, 1, \dots$ )によって表される。ここで、 $f^{(m,s)}(i, j) = (k, l)$ は、始点画像の $p^{(m,s)}_{(i,j)}$ が終点画像の $q^{(m,s)}_{(k,l)}$ に写像されることを意味する。簡単のために、 $f(i, j) = (k, l)$ が成り立つとき画素 $q_{(k,l)}$ を $q_{f(i,j)}$ と\*

$$p^{(m,s)}_{(i,j)} p^{(m,s)}_{(i+1,j)} p^{(m,s)}_{(i+1,j+1)} p^{(m,s)}_{(i,j+1)}$$

(式4)

を考える( $i=0, \dots, 2^m-1, j=0, \dots, 2^m-1$ )。ここでRの各辺(エッジ)の方向を以下のように定める。 ※ 【0035】  
【数5】

$$\overrightarrow{p^{(m,s)}_{(i,j)} p^{(m,s)}_{(i+1,j)}}, \overrightarrow{p^{(m,s)}_{(i+1,j)} p^{(m,s)}_{(i+1,j+1)}}, \overrightarrow{p^{(m,s)}_{(i+1,j+1)} p^{(m,s)}_{(i,j+1)}} \text{ and } \overrightarrow{p^{(m,s)}_{(i,j+1)} p^{(m,s)}_{(i,j)}} \quad (*)$$

(式5)

この正方形は写像fによって終点画像平面における四辺 ★で示される四辺形、形に写像されなければならない。 $f^{(m,s)}(R)$ によつ ★20 【数6】

$$q^{(m,s)}_{f(i,j)} q^{(m,s)}_{f(i+1,j)} q^{(m,s)}_{f(i+1,j+1)} q^{(m,s)}_{f(i,j+1)}$$

(式6)

は、以下の全単射条件を満たす必要がある。

【0036】1. 四辺形 $f^{(m,s)}(R)$ のエッジは互いに交差しない。

2.  $f^{(m,s)}(R)$ のエッジの方向はRのそれらに等しい(図2の場合、時計回り)。

3. 緩和条件として収縮写像(リトラクション: retractions)を許す。

【0037】何らかの緩和条件を設けないかぎり、全単射条件を完全に満たす写像は単位写像しかないためである。ここでは $f^{(m,s)}(R)$ のひとつのエッジの長さが0、すなわち $f^{(m,s)}(R)$ は三角形になってもよい。しかし、面積が0となるような図形、すなわち1点または1本の線分になつてはならない。図2(R)がもとの四辺形の場合、図2(A)と図2(D)は全単射条件を満たすが、図2(B)、図2(C)、図2(E)は満たさない。

【0038】実際のインプリメンテーションでは、写像★40 【数7】

$$C^{(m,s)}_{(i,j)} = |V(p^{(m,s)}_{(i,j)}) - V(q^{(m,s)}_{f(i,j)})|^2 \quad (*)$$

(式7)

ここで、 $V(p^{(m,s)}_{(i,j)})$ 及び $V(q^{(m,s)}_{f(i,j)})$ はそれぞれ画素 $p^{(m,s)}_{(i,j)}$ 及び $q^{(m,s)}_{f(i,j)}$ の輝度である。 $f$ のトータルエネルギー $C^{(m,s)}$ は、マッチングを評価するひとつの評価式であり、つぎに示すC ◆

$$C^{(m,s)}_f = \sum_{i=0}^{2^m-1} \sum_{j=0}^{2^m-1} C^{(m,s)}_{(i,j)}$$

(式8)

\* 記述する。

【0034】実施の形態で扱う画素(格子点)のようにデータが離散の場合、全単射の定義は重要である。ここでは以下のように定義する( $i, i', j, j', k, l$ は全て整数とする)。まず始めに、始点画像の平面においてRによって表記される各正方形領域、  
【数4】

※ 【0035】  
【数5】

※

★で示される四辺形、

★20 【数6】

30

☆が全射であることを容易に保証すべく、さらに以下の条件を課してもよい。つまり始点画像の境界上の各画素は、終点画像において同じ位置を占める画素に写影されるというものである。すなわち、 $f(i, j) = (i, j)$ (ただし $i=0, i=2^m-1, j=0, j=2^m-1$ の4本の線上)である。この条件を以下「付加条件」とも呼ぶ。

【0039】[1. 3. 2] 写像のエネルギー

[1. 3. 2. 1] 画素の輝度に関するコスト

写像fのエネルギーを定義する。エネルギーが最小になる写像を探すことが目的である。エネルギーは主に、始点画像の画素の輝度とそれに対応する終点画像の画素の輝度の差で決まる。すなわち、写像 $f^{(m,s)}$ の点( $i, j$ )におけるエネルギー $C^{(m,s)}_{(i,j)}$ は次式によって定まる。

【0040】

【数7】

◆ $C^{(m,s)}$ の合計で定義できる。

【0041】

【数8】

15 [1.3.2.2] 滑らかな写像のための画素の位置に関するコスト

滑らかな写像を得るために、写像に関する別のエネルギー $D_r$ を導入する。このエネルギーは画素の輝度とは関係なく、 $p^{(m,s)}_{(i,j)}$  および  $q^{(m,s)}_{(i,j)}$  の位置によ

$$D_{(i,j)}^{(m,s)} = \eta E_{0(i,j)}^{(m,s)} + E_{1(i,j)}^{(m,s)}$$

(式9)

ただし、係数パラメータ $\eta$ は0以上の実数であり、ま

た、

$$E_{0(i,j)}^{(m,s)} = \|(i,j) - f^{(m,s)}(i,j)\|^2$$

(式10)

【数11】

$$E_{1(i,j)}^{(m,s)} = \sum_{i'=i-1}^i \sum_{j'=j-1}^j \|(f^{(m,s)}(i,j) - (i,j)) - (f^{(m,s)}(i',j') - (i',j'))\|^2/4$$

(式11)

とする。ここで、

$$\|(x,y)\| = \sqrt{x^2 + y^2}$$

(式12)

であり、 $i' < 0$  および  $j' < 0$  に対して  $f(i', j')$  は0と決める。 $E_0$  は  $(i, j)$  及び  $f(i, j)$  の距離で決まる。 $E_0$  は画素があまりにも離れた画素へ写影されることを防ぐ。ただし $E_0$  は、後に別のエネルギー関数で置き換える。 $E_1$  は写像の滑らかさを保証する。 $E_1$  は、 $p_{(i,j)}$  の変位とその隣接点の変位の間☆

$$D_f^{(m,s)} = \sum_{i=0}^{2^m-1} \sum_{j=0}^{2^m-1} D_{(i,j)}^{(m,s)}$$

(式13)

[1.3.2.3] 写像の総エネルギー

写像の総エネルギー、すなわち複数の評価式の統合に係る総合評価式は $\lambda C^{(m,s)} + D^{(m,s)}$ で定義される。ここで係数パラメータ $\lambda$ は0以上の実数である。目的は総合評価式が極値をとる状態を検出すること、すなわち次◆

$$\min_f \lambda C_f^{(m,s)} + D_f^{(m,s)}$$

(式14)

$\lambda = 0$  及び  $\eta = 0$  の場合、写像は単位写像になることに注意すべきである（すなわち、全ての  $i = 0, \dots, 2^m - 1$  及び  $j = 0, \dots, 2^m - 1$  に対して  $f^{(m,s)}(i, j) = (i, j)$  となる）。後述のごとく、本実施の形態では最初に  $\lambda = 0$  及び  $\eta = 0$  の場合を評価するため、写像を単位写像から徐々に変形していくことができる。仮に総合評価式の $\lambda$ の位置を変えて  $C^{(m,s)} + \lambda D^{(m,s)}$  と定義したとすれば、 $\lambda = 0$  及び  $\eta = 0$  の場合に総合評価式が  $C^{(m,s)}$  だけになり、本来何等関連のない画素どうしが単に輝度が近いというだけで対応づけら

16 \* って決まる ( $i = 0, \dots, 2^m - 1, j = 0, \dots, 2^m - 1$ )。点  $(i, j)$  における写像  $f^{(m,s)}$  のエネルギー  $D^{(m,s)}_{(i,j)}$  は次式で定義される。

【0042】

【数9】

※【数10】

※10

★ ★【数12】

☆の隔たりを表す。以上の考察をもとに、マッチングを評価する別の評価式であるエネルギー $D_r$ は次式で定まる。

【0043】

【数13】

◆式で示す最小エネルギーを与える写像を見いだすことである。

【0044】

【数14】

れ、写像が無意味なものになる。そうした無意味な写像をもとに写像を変形していてもまったく意味をなさない。このため、単位写像が評価の開始時点で最良の写像として選択されるよう係数パラメータの与えかたが配慮されている。

【0045】オブティカルフローもこの実施の形態同様、画素の輝度の差と滑らかさを考慮する。しかし、オブティカルフローは画像の変換に用いることはできない。オブジェクトの局所的な動きしか考慮しないためである。実施の形態に係る特異点フィルタを用いることに



よって大域的な対応関係を検出することができる。

【0046】[1. 3. 3]多重解像度の導入による写像の決定

最小エネルギーを与え、全単射条件を満足する写像  $f_{m,1}$  を多重解像度の階層を用いて求める。各解像度レベルにおいて始点副画像及び終点副画像間の写像を計算する。解像度の階層の最上位（最も粗いレベル）からスタートし、各解像度レベルの写像を、他のレベルの写像を\*

$$(i', j') = ([\frac{i}{2}], [\frac{j}{2}])$$

が成り立つとき、 $p^{(m-1,*)}(\dots)$ 、 $q^{(m-1,*)}(\dots)$  をそれぞれ  $p^{(*)}(\dots)$ 、 $q^{(*)}(\dots)$  の parent と呼ぶことにする。[x] は x を越えない最大整数である。また  $p^{(*)}(\dots)$ 、 $q^{(*)}(\dots)$  をそれぞれ p ※

$$parent(i, j) = ([\frac{i}{2}], [\frac{j}{2}])$$

$p^{(*)}(\dots)$  と  $q^{(*)}(\dots)$  の間の写像  $f^{(*)}$  は、エネルギー計算を行って最小になったものを見つけることで決定される。 $f^{(*)}(i, j) = (k, 1)$  の値は 20  $f^{(*)}(m=1, 2, \dots, n)$  を用いることによつて、以下のように決定される。まず、 $q^{(*)}(\dots)$  は次★

$$q_{g^{(m,s)}(i-1,j-1)}^{(m,s)} q_{g^{(m,s)}(i-1,j+1)}^{(m,s)} q_{g^{(m,s)}(i+1,j+1)}^{(m,s)} q_{g^{(m,s)}(i+1,j-1)}^{(m,s)}$$

(式 17)

ただしここで、

☆ ☆ 【数18】

$$g^{(m,s)}(i, j) = f^{(m-1,s)}(parent(i, j)) + f^{(m-1,s)}(parent((i, j) + (1, 1)))$$

(式 18)

である。こうして定めた四辺形を、以下  $p^{(*)}(\dots)$  の相続 (inherited) 四辺形と呼ぶことにする。相続四辺形の内部において、エネルギーを最小にする画素を求める。

【0050】図3は以上の手順を示している。同図において、始点画像のA、B、C、Dの画素は、第m-1レベルにおいてそれぞれ終点画像のA'、B'、C'、D'へ写影される。画素  $p^{(*)}(\dots)$  は、相続四辺形A' B' C' D' の内部に存在する画素  $q^{(*)}$  ◆

$$E_{0(i,j)} = \|f^{(m,0)}(i, j) - g^{(m)}(i, j)\|^2$$

(式 19)

また、副写像  $f^{(*)}$  を計算するためには次式を用いる。

\* 【0053】

\* 【数20】

$$E_{0(i,j)} = \|f^{(m,s)}(i, j) - f^{(m,s-1)}(i, j)\|^2 \quad (1 \leq i)$$

(式 20)

こうしてすべての副写像のエネルギーを低い値に保つ写像が得られる。式20により、異なる特異点に対応する

50

\*考慮に入れながら決定する。各レベルにおける写像の候補の数は、より高い、つまりより粗いレベルの写像を用いることによって制限される。より具体的には、あるレベルにおける写像の決定に際し、それよりひとつ粗いレベルにおいて求められた写像が一種の拘束条件として課される。

【0047】まず、

【数15】

(式 15)

※  $p^{(*)}(\dots)$ 、 $q^{(*)}(\dots)$  の child と呼ぶ。関数  $parent(i, j)$  は次式で定義される。

【0048】

【数16】

(式 16)

★の四辺形の内部になければならないという条件を課し、全単射条件を満たす写像のうち現実性の高いものを絞り込む。

【0049】

【数17】

◆  $r^{(*)}(\dots)$  へ写影されなければならない。以上の配慮により、第m-1レベルの写像から第mレベルの写像への橋渡しがなされる。

【0051】先に定義したエネルギーE<sub>0</sub>は、第mレベルにおける副写像  $f^{(*)}$  を計算するために、次式に置き換える。

【0052】

【数19】

副写像が、副写像どうしの類似度が高くなるように同一レベル内で関連づけられる。式19は、 $f^{(*)}(i,$

j)と、第m-1レベルの画素の一部と考えた場合の(i, j)が射影されるべき点の位置との距離を示している。

【0054】仮に、相繞四辺形A' B' C' D'の内部に全単射条件を満たす画素が存在しない場合は以下の措置をとる。まず、A' B' C' D'の境界線からの距離がL(始めはL=1)である画素を調べる。それらのうち、エネルギーが最小になるものが全単射条件を満たせば、これを $f^{(m)}(i, j)$ の値として選択する。そのような点が発見されるか、またはLがその上限の $L^{(m)}_{max}$ に到達するまで、Lを大きくしていく。L<sup>(m)</sup><sub>max</sub>は各レベルmに対して固定である。そのような点が全く発見されない場合、全単射の第3の条件を一時的に無視して変換先の四辺形の面積がゼロになるような写像も認め、 $f^{(m)}(i, j)$ を決定する。それでも条件を満たす点が見つからない場合、つぎに全単射の第1及び第2条件を外す。

【0055】多重解像度を用いる近似法は、写像が画像の細部に影響されることを回避しつつ、画像間の大域的な対応関係を決定するために必須である。多重解像度による近似法を用いなければ、距離の遠い画素間の対応関係を見いだすことは不可能である。その場合、画像のサイズはきわめて小さなものに限定しなければならず、変化の小さな画像しか扱うことができない。さらに、通常写像に滑らかさを要求するため、そうした画素間の対応関係を見つけにくくしている。距離のある画素から画素への写像のエネルギーは高いためである。多重解像度を用いた近似法によれば、そうした画素間の適切な対応関係を見いだすことができる。それらの距離は、解像度の階層の上位レベル(粗いレベル)において小さいためである。

【0056】[1.4]最適なパラメータ値の自動決定  
既存のマッチング技術の主な欠点のひとつに、パラメータ調整の困難さがある。大抵の場合、パラメータの調整は人手作業によって行われ、最適な値を選択することはきわめて難しい。実施の形態に係る方法によれば、最適なパラメータ値を完全に自動決定することができる。

【0057】実施の形態に係るシステムはふたつのパラメータ、λ及びηを含む。端的に言えば、λは画素の輝度の差の重みであり、ηは写像の剛性を示している。これらのパラメータの値は初期値が0であり、まずη=0に固定してλを0から徐々に増加させる。λの値を大きくしながら、しかも総合評価式(式14)の値を最小にする場合、各副写像に関する $C^{(m)}$ の値は一般に小さく

＊くなっていく。このことは基本的にふたつの画像がよりマッチしなければならないことを意味する。しかし、λが最適値を超えると以下の現象が発生する。

【0058】1. 本来対応すべきではない画素どうしが、単に輝度が近いというだけで誤って対応づけられる。

2. その結果、画素どうしの対応関係がおかしくなり、写像がくずれはじめる。

【0059】3. その結果、式14において $D^{(m)}$ が急激に増加しようとする。

4. その結果、式14の値が急激に増加しようとするため、 $D^{(m)}$ の急激な増加を抑制するよう $f^{(m)}$ が変化し、その結果 $C^{(m)}$ が増加する。したがって、λを増加させながら式14が最小値をとるという状態を維持しつつ $C^{(m)}$ が減少から増加に転じる閾値を検出し、そのλをη=0における最適値とする。つぎにηを少しずつ増やして $C^{(m)}$ の挙動を検査し、後述の方法でηを自動決定する。そのηに対応してλも決まる。

【0060】この方法は、人間の視覚システムの焦点機構の動作に似ている。人間の視覚システムでは、一方の目を動かしながら左右両目の画像のマッチングがとられる。オブジェクトがはっきりと認識できるとき、その目が固定される。

【0061】[1.4.1]λの動的決定

λは0から所定の刻み幅で増加されていき、λの値が変わる度に副写像が評価される。式14のごとく、総エネルギーは $\lambda C^{(m)} + D^{(m)}$ によって定義される。式9の $D^{(m)}$ は滑らかさを表すもので、理論的には単位写像の場合に最小になり、写像が歪むほど $E_0$ も増加していく。 $E_1$ は整数であるから、 $D^{(m)}$ の最小刻み幅は1である。このため、現在の $\lambda C^{(m)}_{(i,j)}$ の変化(減少量)が1以上でなければ、写像を変化させることによって総エネルギーを減らすことはできない。なぜなら、写像の変化に伴って $D^{(m)}$ は1以上増加するため、 $\lambda C^{(m)}_{(i,j)}$ が1以上減少しない限り総エネルギーは減らないためである。

【0062】この条件のもと、λの増加に伴い、正常な場合に $C^{(m)}_{(i,j)}$ が減少することを示す。 $C^{(m)}_{(i,j)}$ のヒストグラムをh(1)と記述する。h(1)はエネルギー $C^{(m)}_{(i,j)}$ が1<sup>2</sup>である画素の数である。λ1<sup>2</sup>≥1が成り立つために、例えば1<sup>2</sup>=1/λの場合を考える。λがλ<sub>1</sub>からλ<sub>2</sub>まで微小量変化するとき、

【数21】

$$A = \sum_{l=\lfloor \frac{1}{\lambda_2} \rfloor}^{\lfloor \frac{1}{\lambda_1} \rfloor} h(l) \simeq \int_{\frac{1}{\lambda_2}}^{\frac{1}{\lambda_1}} h(l) dl = - \int_{\lambda_2}^{\lambda_1} h(l) \frac{1}{\lambda^{3/2}} d\lambda = \int_{\lambda_1}^{\lambda_2} \frac{h(l)}{\lambda^{3/2}} d\lambda$$

(式21)

で示されるA個の画素が、

$$C_f^{(m,s)} - l^2 = C_f^{(m,s)} - \frac{1}{\lambda} \quad (式22)$$

のエネルギーを持つより安定的な状態に変化する。ここでは仮に、これらの画素のエネルギーがすべてゼロにな

※と近似している。この式はC<sup>(m,s)</sup>の値が、  
【数23】

$$\partial C_f^{(m,s)} = -\frac{A}{\lambda} \quad (式23)$$

だけ変化することを示し、その結果、

★ ★【数24】

$$\frac{\partial C_f^{(m,s)}}{\partial \lambda} = -\frac{h(l)}{\lambda^{5/2}} \quad (式24)$$

が成立する。h(1) > 0であるから、通常C<sup>(m,s)</sup>は減少する。しかし、λが最適値を越えようとするとき、上述の現象、つまりC<sup>(m,s)</sup>の増加が発生する。この現象を検出することにより、λの最適値を決定する。 ☆

☆【0063】なお、H(h > 0)及びkを定数とすると  
【数25】

$$h(l) = H l^k = \frac{H}{\lambda^{k/2}} \quad (式25)$$

と仮定すれば、

◆ ◆【数26】

$$\frac{\partial C_f^{(m,s)}}{\partial \lambda} = -\frac{H}{\lambda^{5/2+k/2}} \quad (式26)$$

が成り立つ。このときk ≠ -3であれば、

※ ※【数27】

$$C_f^{(m,s)} = C + \frac{H}{(3/2 + k/2)\lambda^{3/2+k/2}} \quad (式27)$$

となる。これがC<sup>(m,s)</sup>の一般式である(Cは定数)。

※の率で増加する。

【0064】λの最適値を検出する際、さらに安全を見て、全単射条件を破る画素の数を検査してもよい。ここで各画素の写像を決定する際、全単射条件を破る確率をp<sub>0</sub>と仮定する。この場合、

【0065】

30 【数29】

$$B_0 = \frac{h(l)p_0}{\lambda^{3/2}}$$

【数28】

$$\frac{\partial A}{\partial \lambda} = \frac{h(l)}{\lambda^{3/2}} \quad (式28)$$

従って、

【数30】

が成立しているため、全単射条件を破る画素の数は次式※

$$\frac{B_0 \lambda^{3/2}}{p_0 h(l)} = 1 \quad (式30)$$

は定数である。仮にh(1) = H l<sup>k</sup>を仮定するとき、例えば、

★【数31】

★

$$B_0 \lambda^{3/2+k/2} = p_0 H \quad (式31)$$

は定数になる。しかしλが最適値を越えると、上の値は急速に増加する。この現象を検出し、B<sub>0</sub>λ<sup>3/2+k/2</sup>/2<sup>k</sup>の値が異常値B<sub>0,1,2,3,4,5,6,7,8,9,10</sub>を越えるかどうかを検査し、λの最適値を決定することができる。同様に、B<sub>1</sub>λ<sup>3/2+k/2</sup>/2<sup>k</sup>の値が異常値B<sub>1,1,2,3,4,5,6,7,8,9,10</sub>を越えるかどうかを検査することにより、全単射の第3の条件を破る画素の

増加率B<sub>1</sub>を確認する。ファクター2<sup>k</sup>を導入する理由は後述する。このシステムはこれら2つの閾値に敏感ではない。これらの閾値は、エネルギーC<sup>(m,s)</sup>の観察では検出し損なった写像の過度の歪みを検出するために用いることができる。

【0066】なお実験では、副写像f<sup>(m,s)</sup>を計算する

際、もし $\lambda$ が0.1を超えたら $f^{(n)}$ の計算は止めて $f^{(n+1)}$ の計算に移行した。 $\lambda > 0.1$ のとき、画素の輝度255レベル中のわずかに「3」の違いが副写像の計算に影響したためであり、 $\lambda > 0.1$ のとき正しい結果を得ることは困難だったためである。

【0067】[1.4.2]ヒストグラム $h(1)$   
 $C^{(n)}$ の検査はヒストグラム $h(1)$ に依存しない。全単射及びその第3の条件の検査の際、 $h(1)$ に影響を受けうる。実際に $(\lambda, C^{(n)})$ をプロットすると、 $k$ は通常1付近にある。実験では $k=1$ を用い、 $B$  10  
 $\lambda^2$ と $B_1 \lambda^2$ を検査した。仮に $k$ の本当の値が1未満で\*

$$p(i,j) = \begin{cases} \frac{255}{r} c(\sqrt{(i-x_0)^2 + (j-y_0)^2}) & (\sqrt{(i-x_0)^2 + (j-y_0)^2} \leq r) \\ 0 & (\text{otherwise}) \end{cases} \quad (\text{式32})$$

一方、終点画像は、次式のごとく中心 $(x_1, y_1)$ 、半径が $r$ のオブジェクトであるとする。 ※ 【0070】

$$q(i,j) = \begin{cases} \frac{255}{r} c(\sqrt{(i-x_1)^2 + (j-y_1)^2}) & (\sqrt{(i-x_1)^2 + (j-y_1)^2} \leq r) \\ 0 & (\text{otherwise}) \end{cases} \quad (\text{式33})$$

ここで $c(x)$ は $c(x) = x^k$ の形であるとする。中心 $(x_0, y_0)$ 及び $(x_1, y_1)$ が十分遠い場合、ヒストグラム $h(1)$ は次式の形となる。

【0071】

【数34】

$$h(i) \propto r i^k \quad (k \neq 0) \quad (\text{式34})$$

$k=1$ のとき、画像は背景に埋め込まれた鮮明な境界線を持つオブジェクトを示す。このオブジェクトは中心が暗く、周囲にいくに従って明るくなる。 $k=-1$ のとき、画像は曖昧な境界線を持つオブジェクトを表す。このオブジェクトは中心が最も明るく、周囲にいくに従って暗くなる。一般のオブジェクトはこれらふたつのタイプのオブジェクトの中間にあると考えてもさして一般性を失わない。したがって、 $k$ は $-1 \leq k \leq 1$ として大抵の場合をカバーでき、式27が一般に減少関数であること★

$$E_{0(i,j)}^{(m,s)} = \|f^{(m,s)}(i,j) - f^{(m,s-1)}(i,j)\|^2$$

(式35)

$\eta$ が0のとき、 $D^{(n)}$ は直前の副写像と無関係に決定され、現在の副写像は弾性的に変形され、過度に歪むことになる。一方、 $\eta$ が非常に大きな値のとき、 $D^{(n)}$ は直前の副写像によってほぼ完全に決まる。このとき副写像は非常に剛性が高く、画素は同じ場所に射影される。その結果、写像は単位写像になる。 $\eta$ の値が0から次第に増えるとき、後述のごとく $C^{(n)}$ は徐々に減少する。しかし $\eta$ の値が最適値を超えると、図4に示すとおり、エ

\*あれば、 $B_0 \lambda^2$ と $B_1 \lambda^2$ は定数にならず、ファクター $\lambda^{(1-k)/2}$ に従って徐々に増加する。 $h(1)$ が定数であれば、例えばファクターは $\lambda^{1/2}$ である。しかし、こうした差は閾値 $B_{0,1,2,3,4}$ を正しく設定することによって吸収することができる。

【0068】ここで次式のごとく始点画像を中心が $(x_0, y_0)$ 、半径 $r$ の円形のオブジェクトであると仮定する。

【0069】

【数32】

※ 【0070】

※ 【数33】

★とが保障される。

【0072】なお、式34からわかるように、 $r$ は画像の解像度に影響されること、すなわち $r$ は $2^\circ$ に比例することに注意すべきである。このために[1.4.1]においてファクター $2^\circ$ を導入した。

【0073】[1.4.3] $\eta$ の動的決定

パラメータ $\eta$ も同様の方法で自動決定できる。はじめに $\eta=0$ とし、最も細かい解像度における最終的な写像 $f^{(n)}$ 及びエネルギー $C^{(n)}$ を計算する。つづいて、 $\eta$ をある値 $\Delta\eta$ だけ増加させ、再び最も細かい解像度における最終写像 $f^{(n)}$ 及びエネルギー $C^{(n)}$ を計算し直す。この過程を最適値が求まるまで続ける。 $\eta$ は写像の剛性を示す。次式の重みだからである。

【0074】

【数35】

エネルギーは増加し始める。同図のX軸は $\eta$ 、Y軸は $C$ である。

【0075】この方法で $C^{(n)}$ を最小にする最適な $\eta$ の値を得ることができる。しかし、 $\lambda$ の場合に比べていろいろな要素が計算に影響する結果、 $C^{(n)}$ は小さく揺らぎながら変化する。 $\lambda$ の場合は、入力が微小量変化するたびに副写像を1回計算しなおすだけだが、 $\eta$ の場合はすべての副写像が計算しなおされるためである。このた

25

め、得られた $C^{(n)}$ の値が最小であるかどうかを即座に判断することはできない。最小値の候補が見つければ、さらに細かい区間を設定することによって真の最小値を探す必要がある。

【0076】[1.5]スーパーサンプリング

$$V(q_{f(m,s)}^{(m,s)}(i,j))$$

(式36)

における輝度を持つ $f^{(m,s)}$ が提供される。つまりスーパーサンプリングが行われる。実験では、 $f^{(m,s)}$ は整

$$V(q_{(i,j)+(0.5,0.5)}^{(m,s)})$$

(式37)

は、

★ ★【数38】

$$(V(q_{(i,j)}^{(m,s)}) + V(q_{(i,j)+(1,1)}^{(m,s)}))/2$$

(式38)

によって与えられた。

【0077】[1.6]各画像の画素の輝度の正規化  
始点画像と終点画像がきわめて異なるオブジェクトを含んでいるとき、写像の計算に元の画素の輝度がそのままでは利用しにくい。輝度の差が大きいために輝度に関するエネルギー $C^{(m,s)}$ が大きくなりすぎ、正しい評価がしずらいためである。

【0078】例えば、図20(a)と図20(b)に示すように人の顔と猫の顔のマッチングをとる場合を考える。猫の顔は毛で覆われており、非常に明るい画素と非常に暗い画素が混じっている。この場合、ふたつの顔の間の副写像を計算するために、まず副画像を正規化する。すなわち、最も暗い画素の輝度を0、最も明るいそれを255に設定し、他の画素の輝度は線形補間によって求めておく。

【0079】[1.7]インプリメンテーション  
始点画像のスキャンに従って計算がリニアに進行する帰納的な方法を用いる。始めに、1番上の左端の画素 $(i,j) = (0,0)$ について $f^{(m,s)}$ の値を決定する。次に $i$ を1ずつ増やしながら各 $f^{(m,s)}(i,j)$ の値を決定する。 $i$ の値が画像の幅に到達したとき、 $j$ の値を1増やし、 $i$ を0に戻す。以降、始点画像のスキャンに伴い $f^{(m,s)}(i,j)$ を決定していく。すべての点について画素の対応が決まれば、ひとつの写像 $f^{(m,s)}$ が決まる。

【0080】ある $p_{r(i,j)}$ について対応点 $q_{r(i,j)}$ が決まれば、つぎに $p_{r(i,j)}$ の対応点 $q_{r(i,j)}$ が決める。この際、 $q_{r(i,j)}$ の位置は全単射条件を満たすために、 $q_{r(i,j)}$ の位置によって制限される。したがって、先に対応点が決まる点ほどこのシステムでは優先度が高くなる。つねに $(0,0)$ が最も優先される状態が★

$$W = \vec{A} \times \vec{B}$$

(式39)

26

\* 画素間の対応関係を決定する際、自由度を増やすために、 $f^{(m,s)}$ の値域を $R \times R$ に拡張することができる( $R$ は実数の集合)。この場合、終点画像の画素の輝度が補間され、非整数点、

\* 【数36】

\* 数及び半整数値をとることが許され、

※10 【数37】

☆ つづくと、求められる最終の写像に余計な偏向が加わる。本実施の形態ではこの状態を回避するために、 $f^{(m,s)}$ を以下の方法で決めていく。

【0081】まず $(s \bmod 4)$ が0の場合、 $(0,0)$ を開始点とし $i$ 及び $j$ を徐々に増やしながら決めていく。 $(s \bmod 4)$ が1の場合、最上行の右端点を開始点とし、 $i$ を減少、 $j$ を増加させながら決めていく。 $(s \bmod 4)$ が2のとき、最下行の右端点を開始点とし、 $i$ 及び $j$ を減少させながら決めていく。 $(s \bmod 4)$ が3の場合、最下行の左端点を開始点とし、 $i$ を増加、 $j$ を減少させながら決めていく。解像度が最も細かい第 $n$ レベルには副写像という概念、すなわちパラメータ $s$ が存在しないため、仮に $s=0$ 及び $s=2$ であるとしてふたつの方向を連続的に計算した。

【0082】実際のインプリメンテーションでは、全単射条件を破る候補に対してペナルティを与えることにより、候補 $(k,1)$ の中からできる限り全単射条件を満たす $f^{(m,s)}(i,j)$  ( $m=0, \dots, n$ )の値を選んだ。第3の条件を破る候補のエネルギー $D(k,1)$ には $\psi$ を掛け、一方、第1または第2の条件を破る候補には $\phi$ を掛ける。今回は $\phi=2$ 、 $\psi=100000$ を用いた。

【0083】前述の全単射条件のチェックのために、実際の手続として $(k,1) = f^{(m,s)}(i,j)$ を決定する際に以下のテストを行った。すなわち $f^{(m,s)}(i,j)$ の相続四辺形に含まれる各格子点 $(k,1)$ に対し、次式の外積の $z$ 成分が0以上になるかどうかを確かめる。

【0084】

【数39】

ただしここで、

\* \* 【数40】

$$\vec{A} = \overrightarrow{q_{f(m,s)}^{(m,s)}(i,j-1) q_{f(m,s)}^{(m,s)}(i+1,j-1)}$$

(式40)

【数41】

$$\vec{B} = \overrightarrow{q_{f(m,s)}^{(m,s)}(i,j-1) q_{(k,l)}^{(m,s)}}$$

(式41)

である(ここでベクトルは三次元ベクトルとし、z軸は直交右手座標系において定義される)。もしWが負であれば、その候補については $D^{(m,s)}(i,j)$ にψを掛けることによってペナルティを与え、できるかぎり選択しないようにする。

【0085】図5(a)、図5(b)はこの条件を検査する理由を示している。図5(a)はペナルティのない候補、図5(b)はペナルティがある候補をそれぞれ表す。隣接画素(i, j+1)に対する写像 $f^{(m,s)}(i, j+1)$ を決定する際、Wのz成分が負であれば始点画像平面上において全単射条件を満足する画素は存在しない。なぜなら、 $q^{(m,s)}(i,j)$ は隣接する四辺形の境界線を越えるためである。

【0086】【1. 7. 1】副写像の順序  
インプリメンテーションでは、解像度レベルが偶数のときは $\sigma(0)=0, \sigma(1)=1, \sigma(2)=2, \sigma(3)=3, \sigma(4)=0$ を用い、奇数のときは $\sigma(0)=3, \sigma(1)=2, \sigma(2)=1, \sigma(3)=0$ を用い、

$$\begin{aligned} (x,y) = & (1-dx)(1-dy)(1-t)(i,j) + (1-dx)(1-dy)tf(i,j) \\ & + dx(1-dy)(1-t)(i+1,j) + dx(1-dy)tf(i+1,j) \\ & + (1-dx)dy(1-t)(i,j+1) + (1-dx)dytf(i,j+1) \\ & + dx dy(1-t)(i+1,j+1) + dx dy tf(i+1,j+1) \end{aligned}$$

(式42)

つづいて $r(x, y, t)$ における画素の輝度が次の式 ★【0089】  
を用いて決定される。 ★ 【数43】

$$\begin{aligned} V(r(x,y,t)) = & (1-dx)(1-dy)(1-t)V(p(i,j)) + (1-dx)(1-dy)tfV(q_f(i,j)) \\ & + dx(1-dy)(1-t)V(p(i+1,j)) + dx(1-dy)tfV(q_f(i+1,j)) \\ & + (1-dx)dy(1-t)V(p(i,j+1)) + (1-dx)dytfV(q_f(i,j+1)) \\ & + dx dy(1-t)V(p(i+1,j+1)) + dx dy tfV(q_f(i+1,j+1)) \end{aligned}$$

(式43)

ここでdx及びdyはパラメータであり、0から1まで変化する。

【0090】【1. 9】拘束条件を課したときの写像

※0、 $\sigma(4)=3$ を用いた。このことで、副写像を適度にシャッフルした。なお、本来副写像は4種類であり、sは0~3のいずれかである。しかし、実際には $s=4$ に相当する処理を行った。その理由は後述する。

【0087】【1. 8】補間計算

始点画像と終点画像の間の写像が決定された後、対応しあう画素の輝度が補間される。実験では、トライリニア補間を用いた。始点画像平面上における正方形 $p(i,j), p(i+1,j), p(i,j+1), p(i+1,j+1)$ が終点画像平面上の四辺形 $q_f(i,j), q_f(i+1,j), q_f(i,j+1), q_f(i+1,j+1)$ に射影されると仮定する。簡単のため、画像間の距離を1とする。始点画像平面からの距離がt( $0 \leq t \leq 1$ )である中間画像の画素 $r(x, y, t)$ ( $0 \leq x \leq N-1, 0 \leq y \leq M-1$ )は以下の要領で求められる。まず画素 $r(x, y, t)$ の位置(ただし $x, y, t \in \mathbb{R}$ )を次式で求める。

【0088】

【数42】

いままでは拘束条件がいっさい存在しない場合の写像の決定を述べた。しかし、始点画像と終点画像の特定の画素間に予め対応関係が規定されているとき、これを拘束

条件としたうえで写像を決定することができる。

【0091】基本的な考えは、まず始点画像の特定の画素を終点画像の特定の画素に移す大まかな写像によって始点画像を大まかに変形し、しかる後、写像  $f$  を正確に計算する。

【0092】まず始めに、始点画像の特定の画素を終点画像の特定の画素に射影し、始点画像の他の画素を適当な位置に射影する大まかな写像を決める。すなわち、特\*

$$P(i_0, j_0), P(i_1, j_1), \dots,$$

\* 定の画素に近い画素は、その特定の画素が射影される場所の近くに射影されるような写像である。ここで第  $m$  レベルの大まかな写像を  $F^{(m)}$  と記述する。

【0093】大まかな写像  $F$  は以下の要領で決める。まず、いくつかの画素について写像を特定する。始点画像について  $n_s$  個の画素、

【数44】

$$P(i_{n_s-1}, j_{n_s-1})$$

(式44)

を特定するとき、以下の値を決める。

※【数45】

【0094】

※

$$\begin{aligned} F^{(n)}(i_0, j_0) &= (k_0, l_0), \\ F^{(n)}(i_1, j_1) &= (k_1, l_1), \dots, \end{aligned}$$

$$F^{(n)}(i_{n_s-1}, j_{n_s-1}) = (k_{n_s-1}, l_{n_s-1})$$

(式45)

始点画像の他の画素の変位量は、 $P(i_h, j_h)$  ( $h=0, \dots, n_s-1$ ) の変位に重み付けをして求められる平均である。すなわち画素  $P(i, j)$  は、終点画像の以下の画 ★

★素に射影される。

【0095】

【数46】

$$F^{(m)}(i, j) = \frac{(i, j) + \sum_{h=0}^{n_s-1} (k_h - i_h, l_h - j_h) \text{weight}_h(i, j)}{2^{n-m}}$$

(式46)

ただしここで、

☆ ☆【数47】

$$\text{weight}_h(i, j) = \frac{1/\|(i_h - i, j_h - j)\|^2}{\text{total weight}(i, j)}$$

(式47)

【数48】

$$\text{total weight}(i, j) = \sum_{h=0}^{n_s-1} 1/\|(i_h - i, j_h - j)\|^2$$

(式48)

とする。

【0096】つづいて、 $F^{(m)}$  に近い候補写像  $f$  がより

少ないエネルギーを持つように、その写像  $f$  のエネルギー ◆

◆  $D^{(m,s)}(i, j)$  を変更する。正確には、 $D^{(m,s)}(i, j)$  は、

【数49】

$$D_{(i,j)}^{(m,s)} = E_{0(i,j)}^{(m,s)} + \eta E_{1(i,j)}^{(m,s)} + \kappa E_{2(i,j)}^{(m,s)}$$

(式49)

である。ただし、

【数50】

$$E_{2(i,j)}^{(m,s)} = \begin{cases} 0, & \text{if } \|F^{(m)}(i,j) - f^{(m,s)}(i,j)\|^2 \leq \lfloor \frac{\rho^2}{2^{2(n-m)}} \rfloor \\ \|F^{(m)}(i,j) - f^{(m,s)}(i,j)\|^2, & \text{otherwise} \end{cases}$$

(式50)

であり、 $\kappa, \rho \geq 0$ とする。最後に、前述の写像の自動計算プロセスにより、 $f$ を完全に決定する。

【0097】ここで、 $f^{(m,s)}(i,j)$ が $F^{(m)}(i,j)$ に十分近いとき、つまりそれらの距離が、  
【数51】

$$\left\lfloor \frac{\rho^2}{2^{2(n-m)}} \right\rfloor \quad (\text{式51})$$

以内であるとき、 $E_{2(i,j)}^{(m,s)}$ が0になることに注意すべきである。そのように定義した理由は、各 $f^{(m,s)}(i,j)$ が $F^{(m)}(i,j)$ に十分近い限り、終点画像において適切な位置に落ち着くよう、その値を自動的に決めたいためである。この理由により、正確な対応関係を詳細に特定する必要がなく、始点画像は終点画像にマッチするように自動的にマッピングされる。

【0098】[2]具体的な処理手順

[1]の各要素技術による処理の流れを説明する。図6は実施の形態の全体手順を示すフローチャートである。同図のごとく、まず多重解像度特異点フィルタを用いた処理を行い(S1)、つづいて始点画像と終点画像のマッチングをとる(S2)。ただし、S2は必須ではなく、S1で得られた画像の特徴をもとに画像認識などの処理を行ってもよい。

【0099】図7は図6のS1の詳細を示すフローチャートである。ここではS2で始点画像と終点画像のマッチングをとることを前提としている。そのため、まず特異点フィルタによって始点画像の階層化を行い(S10)、一連の始点階層画像を得る。つづいて同様の方法で終点画像の階層化を行い(S11)、一連の終点階層画像を得る。ただし、S10とS11の順序は任意であるし、始点階層画像と終点階層画像を並行して生成していくこともできる。

【0100】図8は図7のS10の詳細を示すフローチャートである。もとの始点画像のサイズは $2^n \times 2^n$ とする。始点階層画像は解像度が細かいほうから順に作られるため、処理の対象となる解像度レベルを示すパラメータ $m$ を $n$ にセットする(S100)。つづいて第 $m$ レベルの画像 $p^{(m,0)}, p^{(m,1)}, p^{(m,2)}, p^{(m,3)}$ から特異点フィルタを用いて特異点を検出し(S101)、それぞれ第 $m-1$ レベルの画像 $p^{(m-1,0)}, p^{(m-1,1)}, p^{(m-1,2)}, p^{(m-1,3)}$ を生成する(S102)。ここでは $m=n$ であるため、 $p^{(n,0)} = p^{(n,1)} = p^{(n,2)} = p^{(n,3)} = p^{(n)}$ であり、ひとつの始点画像から4種類の副画像が生成される。

【0101】図9は第 $m$ レベルの画像の一部と、第 $m-1$ レベルの画像の一部の対応関係を示している。同図の

数値は各画素の輝度を示す。同図の $p^{(m,s)}$ は $p^{(m,0)} \sim p^{(m,3)}$ の4つの画像を象徴するもので、 $p^{(m-1,0)}$ を生成する場合には、 $p^{(m,s)}$ は $p^{(m,0)}$ であると考える。

[1,2]で示した規則により、 $p^{(m-1,0)}$ は例えば同図で輝度を記入したブロックについて、そこに含まれる4画素のうち「3」、 $p^{(m-1,1)}$ は「8」、 $p^{(m-1,2)}$ は「6」、 $p^{(m-1,3)}$ を「10」をそれぞれ取得し、このブロックをそれぞれ取得したひとつの画素で置き換える。したがって、第 $m-1$ レベルの副画像のサイズは $2^{m-1} \times 2^{m-1}$ になる。

【0102】つづいて $m$ をデクリメントし(図8のS103)、 $m$ が負になっていないことを確認し(S104)、S101に戻ってつぎに解像度の粗い副画像を生成していく。この繰り返し処理の結果、 $m=0$ 、すなわち第0レベルの副画像が生成された時点でS10が終了する。第0レベルの副画像のサイズは $1 \times 1$ である。

【0103】図10はS10によって生成された始点階層画像を $n=3$ の場合について例示している。最初の始点画像のみが4つの系列に共通であり、以降特異点の種類に応じてそれぞれ独立に副画像が生成されていく。なお、図8の処理は図7のS11にも共通であり、同様の手順を経て終点階層画像も生成される。以上で図6のS1による処理が完了する。

【0104】実施の形態では、図6のS2に進むためにマッチング評価の準備をする。図11はその手順を示している。同図のごとく、まず複数の評価式が設定される(S30)。[1,3,2,1]で導入した画素に関するエネルギー $C^{(m,s)}$ と[1,3,2,2]で導入した写像の滑らかさに関するエネルギー $D^{(m,s)}$ がそれぞれある。つぎに、これらの評価式を統合して総合評価式を立てる(S31)。[1,3,2,3]で導入した総エネルギー $\lambda C^{(m,s)} + D^{(m,s)}$ がそれぞれあり、[1,3,2,2]で導入した $\eta$ を用いれば、 $\sum (\lambda C^{(m,s)}(i,j) + \eta E_0^{(m,s)}(i,j) + E_1^{(m,s)}(i,j))$  (式52)となる。ただし、総和は $i, j$ についてそれぞれ $0, 1, \dots, 2^{m-1}$ で計算する。以上でマッチング評価の準備が整う。

【0105】図12は図6のS2の詳細を示すフローチャートである。[1]で述べたごとく、始点階層画像と終点階層画像のマッチングは互いに同じ解像度レベルの画像どうしでとられる。画像間の大域的なマッチングを良好にとるために、解像度が粗いレベルから順にマッチングを計算する。特異点フィルタを用いて始点階層画像および終点階層画像を生成しているため、特異点の位置や輝度は解像度の粗いレベルでも明確に保存されており、大域的なマッチングの結果は従来に比べて非常に優



れたものになる。

【0106】図12のごとく、まず係数パラメータ $\eta$ を0、レベルパラメータ $m$ を0に設定する(S20)。つづいて、始点階層画像中の第 $m$ レベルの4つの副画像と終点階層画像中の第 $m$ レベルの4つの副画像のそれぞれの間でマッチングを計算し、それぞれ全単射条件を満たし、かつエネルギーを最小にするような4種類の副写像 $f^{(s)}$  ( $s=0, 1, 2, 3$ )を求める(S21)。

全単射条件は[1. 3. 3]で述べた相続四辺形を用いて検査される。この際、式17、18が示すように、第 $m$ レベルにおける副写像は第 $m-1$ レベルのそれらに拘束されるため、より解像度の粗いレベルにおけるマッチングが順次利用されていく。これは異なるレベル間の垂直的参照である。なお、いま $m=0$ であってそれより粗いレベルはないが、この例外的な処理は図13で後述する。

【0107】一方、同一レベル内における水平的参照も行われる。[1. 3. 3]の式20のごとく、 $f^{(s)}$ は $f^{(s)}$ に、 $f^{(s)}$ は $f^{(s)}$ に、 $f^{(s)}$ は $f^{(s)}$ に、それぞれ類似するように決める。その理由は、特異点の種類が違って、それらがもともと同じ始点画像と終点画像に含まれている以上、副写像がまったく異なるという状況は不自然だからである。式20からわかるように、副写像どうしが近いほどエネルギーは小さくなり、マッチングが良好とみなされる。

【0108】なお、最初に決めるべき $f^{(s)}$ については同一のレベルで参照できる副写像がないため、式19に示すごとくひとつ粗いレベルを参照する。ただし、実験では $f^{(s)}$ まで求めた後、これを拘束条件として $f^{(s)}$ を一回更新するという手続をとった。これは式20に $s=4$ を代入し、 $f^{(s)}$ を新たな $f^{(s)}$ とすることに等しい。 $f^{(s)}$ と $f^{(s)}$ の間隔度が低くなり過ぎる傾向を回避するためであり、この措置によって実験結果がより良好になった。この措置に加え、実験では

[1. 7. 1]に示す副写像のシャッフルも行った。これも本来特異点の種類ごとに決まる副写像どうしの間隔度を密接に保つ趣旨である。また、処理の開始点に依存する傾向を回避するために、 $s$ の値にしたがって開始点の位置を変える点は[1. 7]で述べたとおりである。

【0109】図13は第0レベルにおいて副写像を決定する様子を示す図である。第0レベルでは各副画像がただひとつの画素で構成されるため、4つの副写像 $f^{(s)}$ はすべて自動的に単位写像に決まる。図14は第1レベルにおいて副写像を決定する様子を示す図である。第1レベルでは副画像がそれぞれ4画素で構成される。同図ではこれら4画素が実線で示されている。いま、 $p^{(s)}$ の点 $x$ の対応点を $q^{(s)}$ の中で探すとき、以下の手順を踏む。

【0110】1. 第1レベルの解像度で点 $x$ の左上点 $a$ 、右上点 $b$ 、左下点 $c$ 、右下点 $d$ を求める。

2. 点 $a \sim d$ がひとつ粗いレベル、つまり第0レベルにおいて属する画素を探す。図14の場合、点 $a \sim d$ はそれぞれ画素 $A \sim D$ に属する。ただし、画素 $A \sim C$ は本来存在しない仮想的な画素である。

3. 第0レベルですでに求まっている画素 $A \sim D$ の対応点 $A' \sim D'$ を $q^{(s)}$ の中にプロットする。画素 $A' \sim C'$ は仮想的な画素であり、それぞれ画素 $A \sim C$ と同じ位置にあるものとする。

4. 画素 $A$ の中の点 $a$ の対応点 $a'$ が画素 $A'$ の中にあるとみなし、点 $a'$ をプロットする。このとき、点 $a$ が画素 $A$ の中で占める位置(この場合、右下)と、点 $a'$ が画素 $A'$ の中で占める位置が同じであると仮定する。

5. 4と同様の方法で対応点 $b' \sim d'$ をプロットし、点 $a' \sim d'$ で相続四辺形を作る。

6. 相続四辺形の中でエネルギーが最小になるよう、点 $x$ の対応点 $x'$ を探す。対応点 $x'$ の候補として、例えば画素の中心が相続四辺形に含まれるものに限定してもよい。図14の場合、4つの画素がすべて候補になる。

【0111】以上がある点 $x$ の対応点の決定手順である。同様の処理を他のすべての点について行い、副写像を決める。第2レベル以上のレベルでは、次第に相続四辺形の形が崩れていくと考えられるため、図3に示すように画素 $A' \sim D'$ の間隔が空いていく状況が発生する。

【0112】こうして、ある第 $m$ レベルの4つの副写像が決まれば、 $m$ をインクリメントし(図12のS22)、 $m$ が $n$ を超えていないことを確かめて(S23)、S21に戻る。以下、S21に戻るたびに次第に細かい解像度のレベルの副写像を求め、最後にS21に戻ったときに第 $n$ レベルの写像 $f^{(s)}$ を決める。この写像は $\eta=0$ に関して定まったものであるから、 $f^{(s)}$  ( $\eta=0$ )と書く。

【0113】つぎに異なる $\eta$ に関する写像も求めるべく、 $\eta$ を $\Delta\eta$ だけシフトし、 $m$ をゼロクリアする(S24)。新たな $\eta$ が所定の探索打ち切り値 $\eta_{\dots}$ を超えていないことを確認し(S25)、S21に戻り、今回の $\eta$ に関して写像 $f^{(s)}$  ( $\eta=\Delta\eta$ )を求める。この処理を繰り返し、S21で $f^{(s)}$  ( $\eta=i\Delta\eta$ ) ( $i=0, 1, \dots$ )を求めていく。 $\eta$ が $\eta_{\dots}$ を超えたときS26に進み、後述の方法で最適な $\eta=\eta_{\dots}$ を決定し、 $f^{(s)}$  ( $\eta=\eta_{\dots}$ )を最終的に写像 $f^{(s)}$ とする。

【0114】図15は図12のS21の詳細を示すフローチャートである。このフローチャートにより、ある定まった $\eta$ について、第 $m$ レベルにおける副写像が決まる。副写像を決める際、実施の形態では副写像ごとに最適な $\lambda$ を独立して決める。

【0115】同図のごとく、まず $s$ と $\lambda$ をゼロクリアする(S210)。つぎに、そのときの $\lambda$ について(および暗に $\eta$ について)エネルギーを最小にする副写像 $f^{(s)}$ を求め(S211)、これを $f^{(s)}$  ( $\lambda=0$ )と

10

20

30

40

50

替く、異なる $\lambda$ に関する写像も求めるべく、 $\lambda$ を $\Delta\lambda$ だけシフトし、新たな $\lambda$ が所定の探索打ち切り値 $\lambda_{\dots}$ を超えていないことを確認し(S213)、S211に戻り、以降の繰返し処理で $f^{(i)}$  ( $\lambda = i\Delta\lambda$ ) ( $i = 0, 1, \dots$ )を求める。 $\lambda$ が $\lambda_{\dots}$ を超えたときS214に進み、最適な $\lambda = \lambda_{\dots}$ を決定し、 $f^{(i)}$  ( $\lambda = \lambda_{\dots}$ )を最終的に写像 $f^{(i)}$ とする(S214)。

【0116】つぎに、同一レベルにおける他の副写像を求めるべく、 $\lambda$ をゼロクリアし、 $s$ をインクリメントする(S215)。 $s$ が4を超えていないことを確認し(S216)、S211に戻る。 $s = 4$ になれば上述のごとく $f^{(i)}$ を利用して $f^{(i)}$ を更新し、そのレベルにおける副写像の決定を終了する。

【0117】図16は、ある $m$ と $s$ について $\lambda$ を変えながら求められた $f^{(i)}$  ( $\lambda = i\Delta\lambda$ ) ( $i = 0, 1, \dots$ )に対応するエネルギー $C^{(i)}$ の挙動を示す図である。【1.4】で述べたとおり、 $\lambda$ が増加すると通常 $C^{(i)}$ は減少する。しかし、 $\lambda$ が最適値を超えると $C^{(i)}$ は増加に転じる。そこで本実施の形態では $C^{(i)}$ が極小値をとるときの $\lambda$ を $\lambda_{\dots}$ と決める。同図のように $\lambda > \lambda_{\dots}$ の範囲で再度 $C^{(i)}$ が小さくなっていても、その時点ではすでに写像がくずれていて意味をなさないため、最初の極小点に注目すればよい。 $\lambda_{\dots}$ は副写像ごとに独立して決めていき、最後に $f^{(i)}$ についてもひとつ定まる。

【0118】一方、図17は、 $\eta$ を変えながら求められた $f^{(i)}$  ( $\eta = i\Delta\eta$ ) ( $i = 0, 1, \dots$ )に対応するエネルギー $C^{(i)}$ の挙動を示す図である。ここでも $\eta$ が増加すると通常 $C^{(i)}$ は減少するが、 $\eta$ が最適値を超えると $C^{(i)}$ は増加に転じる。そこで $C^{(i)}$ が極小値をとるときの $\eta$ を $\eta_{\dots}$ と決める。図17は図4の横軸のゼロ付近を拡大した図と考えてよい。 $\eta_{\dots}$ が決まれば $f^{(i)}$ を最終決定することができる。

【0119】以上、本実施の形態によれば種々のメリットが得られる。まずエッジを検出する必要がないため、エッジ検出タイプの従来技術の課題を解消できる。また、画像に含まれるオブジェクトに対する先験的な知識も不要であり、対応点の自動検出が実現する。特異点フィルタによれば、解像度の粗いレベルでも特異点の輝度や位置を維持することができ、オブジェクト認識、特徴抽出、画像マッチングに極めて有利である。その結果、人手作業を大幅に軽減する画像処理システムの構築が可能となる。

【0120】なお、本実施の形態について次のような変形技術も考えられる。

【0121】(1) 実施の形態では始点階層画像と終点階層画像の間でマッチングをとる際にパラメータの自動決定を行ったが、この方法は階層画像間ではなく、通常の2枚の画像間のマッチングをとる場合全般に利用できる。

【0122】たとえば2枚の画像間で、画素の輝度の差に関するエネルギー $E_1$ と画素の位置的なずれに関するエネルギー $E_2$ のふたつを評価式とし、これらの線形和 $E_{\dots} = \alpha E_1 + E_2$ を総合評価式とする。この総合評価式の極値付近に注目して $\alpha$ を自動決定する。つまり、いろいろな $\alpha$ について $E_{\dots}$ が最小になるような写像を求める。それらの写像のうち、 $\alpha$ に関して $E_1$ が極小値をとるときの $\alpha$ を最適パラメータと決める。そのパラメータに対応する写像を最終的に両画像間の最適マッチングとみなす。

【0123】これ以外にも評価式の設定にはいろいろな方法があり、例えば $1/E_1$ と $1/E_2$ のように、評価結果が良好なほど大きな値をとるものを採用してもよい。総合評価式も必ずしも線形和である必要はなく、 $n$ 乗和( $n = 2, 1/2, -1, -2$ など)、多項式、任意の関数などを適宜選択すればよい。

【0124】パラメータも、 $\alpha$ のみ、実施の形態のごとく $\eta$ と $\lambda$ のふたつの場合、それ以上の場合など、いずれでもよい。パラメータが3以上の場合はひとつずつ変化させて決めていく。

【0125】(2) 本実施の形態では、総合評価式の値が最小になるよう写像を決めた後、総合評価式を構成するひとつの評価式である $C^{(i)}$ が極小になる点を検出してパラメータを決定した。しかし、こうした二段回処理の代わりに、状況によっては単に総合評価式の最小値が最小になるようにパラメータを決めても効果的である。その場合、例えば $\alpha E_1 + \beta E_2$ を総合評価式とし、 $\alpha + \beta = 1$ なる拘束条件を設けて各評価式を平等に扱うなどの措置を講じてよい。パラメータの自動決定の本質は、エネルギーが最小になるようにパラメータを決めていく点にあるからである。

【0126】(3) 実施の形態では各解像度レベルで4種類の特異点に関する4種類の副画像を生成した。しかし、当然4種類のうち1、2、3種類を選択的に用いてもよい。例えば、画像中に明るい点がひとつだけ存在する状態であれば、極大点に関する $f^{(i)}$ だけで階層画像を生成しても相応の効果が得られるはずである。その場合、同一レベルで異なる副写像は不要になるため、 $s$ に関する計算量が減る効果がある。

【0127】(4) 本実施の形態では特異点フィルタによってレベルがひとつ進むと画素が $1/4$ になった。例えば $3 \times 3$ で1ブロックとし、その中で特異点を探す構成も可能であり、その場合、レベルがひとつ進むと画素は $1/9$ になる。

【0128】(5) 始点画像と終点画像がカラーの場合、それらをまず白黒画像に変換し、写像を計算する。その結果求められた写像を用いて始点のカラー画像を変換する。それ以外の方法として、RGBの各成分について副写像を計算してもよい。

【0129】[4] 実験の結果

この実施の形態を用いて様々な画像を補間することができる。異なるふたつの視点からの画像を補間すると、中間視点からの画像を生成することができる。これはWWWにおいてきわめて有利である。なぜなら、限られた数の画像から任意の視点画像を生成できるからである。ふたりの人の顔の画像を補間すれば、モーフィングを行うことができる。CTまたはMRIのデータのように三次元オブジェクトの断面データ画像を用いれば、補間の結果ボリュームレンダリングのための三次元オブジェクトの正確な形状を再構築することができる。

【0130】図18(a)、図18(b)、図18(c)は、写像が中間視点画像を生成するために用いられた場合を示している。ここでは右目画像と左目画像が補間された。図18(a)は左目から見た始点画像、図18(b)は右目から見た始点画像、図18(c)は簡単のために[1.8]において $t$ の値を0.5としたときの中間画像をそれぞれ示している。

【0131】図19(a)、図19(b)、図19(c)、図19(d)は、写像を用いて人間の顔のモーフィングを行った場合を示している。ここではふたりの顔に補間した。図19(a)は始点画像、図19(b)は終点画像、図19(c)は終点画像に始点画像を重ね合わせた画像、図19(d)は $t=0.5$ のときの中間画像をそれぞれ示している。

【0132】図20(a)、図20(b)は、人と猫の顔の補間に写像を用いた場合を示している。図20(a)は猫の顔、図20(b)は人の顔と猫の顔のモーフィング画像をそれぞれ示している。人の顔として図19(a)の画像を用いた。[1.6]で説明した程度の正規化はこの例においてのみ用いられている。

【0133】図21(a)、図21(b)、図21(c)は、数多くのオブジェクトを含む画像に対して今回の方法を適応した例を示している。図21(a)は始点画像、図21(b)は終点画像、図21(c)は $t=0.5$ のときの中間画像をそれぞれ示す。

【0134】図22(a)、図22(b)、図22(c)、図22(d)は、MRIによって得られた人間の脳の断面画像を補間するために写像を用いた結果を示している。図22(a)は始点画像、図22(b)は終点画像(上部断面)、図22(c)は $t=0.5$ の場合の中間画像をそれぞれ示している。また図22(d)は四つの断面画像を用いてボリュームレンダリングを行った結果を斜め方向から見た様子を示す。オブジェクトは完全に不透明であり、補間の結果画素が $51(=255 \times 0.2)$ 以上となった画素のみが表示されている。再構築されたオブジェクトは中心付近で垂直にカットされ、その内部が表示されている。

【0135】これらの例において、MRIの画像は $256 \times 256$ 画素、それ以外の画像は全て $512 \times 512$ 画素である。画素の画素度は $0 \sim 255$ のいずれかの値を

とる。[1.3.1]で説明した付加条件は、図21(a)～図21(c)の場合を除き、すべての例で用いられている。これら全ての例において $B_{\text{channel}}=0.003$ 及び $B_{\text{channel}}=0.5$ が用いられ、これらの値を変更する必要は全くなかった。各副画像の画素の画素度は図20(a)、図20(b)の場合のみ、正規化された。

【図面の簡単な説明】

【図1】 図1(a)とは図1(b)は、ふたりの人物の顔に平均化フィルタを施して得られる画像、図1(c)と図1(d)は、ふたりの人物の顔に関して実施の形態で求められる $p^{(1,0)}$ の画像、図1(e)と図1(f)は、ふたりの人物の顔に関して実施の形態で求められる $p^{(1,1)}$ の画像、図1(g)と図1(h)は、ふたりの人物の顔に関して実施の形態で求められる $p^{(1,2)}$ の画像、図1(i)と図1(j)は、ふたりの人物の顔に関して実施の形態で求められる $p^{(1,3)}$ の画像をそれぞれディスプレイ上に表示した中間調画像の写真である。

【図2】 図2(R)はもとの四辺形を示す図、図2(A)、図2(B)、図2(C)、図2(D)、図2(E)はそれぞれ相続四辺形を示す図である。

【図3】 始点画像と終点画像の関係、および第 $m$ レベルと第 $m-1$ レベルの関係を相続四辺形を用いて示す図である。

【図4】 パラメータ $\eta$ とエネルギー $C_r$ の関係を示す図である。

【図5】 図5(a)、図5(b)は、ある点に関する写像が全単射条件を満たすか否かを外積計算から求める様子を示す図である。

【図6】 実施の形態の全体手順を示すフローチャートである。

【図7】 図6のS1の詳細を示すフローチャートである。

【図8】 図7のS10の詳細を示すフローチャートである。

【図9】 第 $m$ レベルの画像の一部と、第 $m-1$ レベルの画像の一部の対応関係を示す図である。

【図10】 実施の形態で生成された始点階層画像を示す図である。

【図11】 図6のS2に進む前に、マッチング評価の準備の手順を示す図である。

【図12】 図6のS2の詳細を示すフローチャートである。

【図13】 第0レベルにおいて副写像を決定する様子を示す図である。

【図14】 第1レベルにおいて副写像を決定する様子を示す図である。

【図15】 図12のS21の詳細を示すフローチャートである。

【図16】 ある $f^{(0,0)}$ について $\lambda$ を変えながら求め

39

られた $f^{(i)}$  ( $\lambda = i \Delta \lambda$ ) に対応するエネルギー $C^{(i)}$ の挙動を示す図である。

【図17】  $\eta$ を変えながら求められた $f^{(i)}$  ( $\eta = i \Delta \eta$ ) ( $i = 0, 1, \dots$ ) に対応するエネルギー $C^{(i)}$ の挙動を示す図である。

【図18】 図18(a)、図18(b)、図18(c)はそれぞれ、あるオブジェクトに関する左目画像、右目画像、実施の形態で生成された補間画像をディスプレイ上に表示した中間調画像の写真である。

【図19】 図19(a)、図19(b)、図19(c)、図19(d)はそれぞれ、ある人物の顔、別の人物の顔、それらの重ね合わせ画像、実施の形態で生成されたモーフィング画像をディスプレイ上に表示した中\*

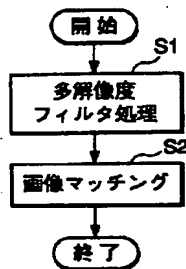
\* 間調画像の写真である。

【図20】 図20(a)、図20(b)はそれぞれ、猫の顔、人と猫の顔のモーフィング画像をディスプレイ上に表示した中間調画像の写真である。

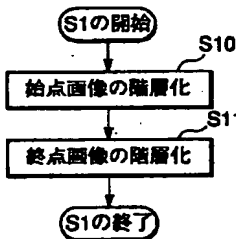
【図21】 図21(a)、図21(b)、図21(c)はそれぞれ、多数のオブジェクトが含まれる左目画像、右目画像、実施の形態で生成された補間画像をディスプレイ上に表示した中間調画像の写真である。

【図22】 図22(a)、図22(b)、図22(c)、図22(d)はそれぞれ、MRに関する始点画像、終点画像、実施の形態で生成された補間画像、補間画像をもとに生成されたボリュームレンダリング画像をディスプレイ上に表示した中間調画像の写真である。

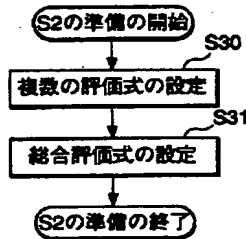
【図6】



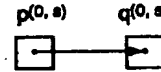
【図7】



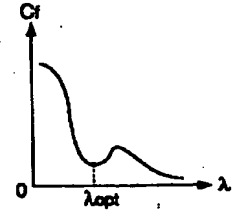
【図11】



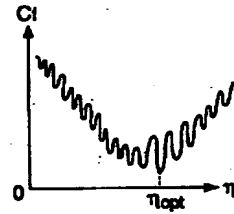
【図13】



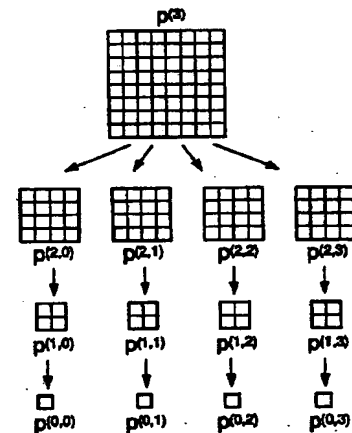
【図16】



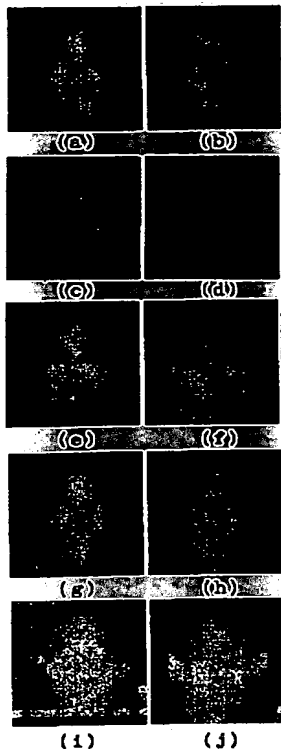
【図17】



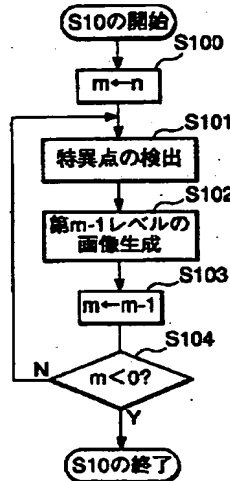
【図10】



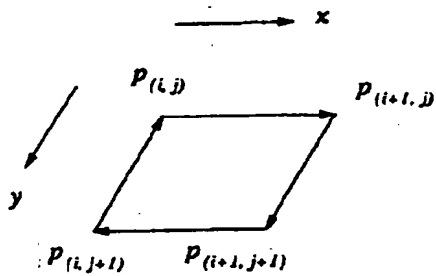
【図1】



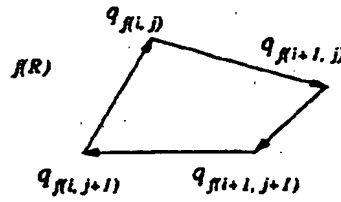
【図8】



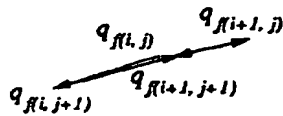
【図2】



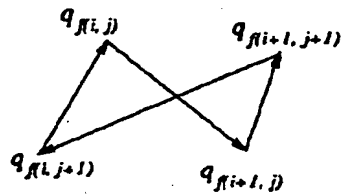
(R)



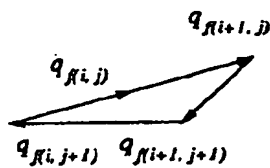
(A)



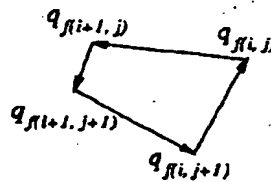
(E)



(B)

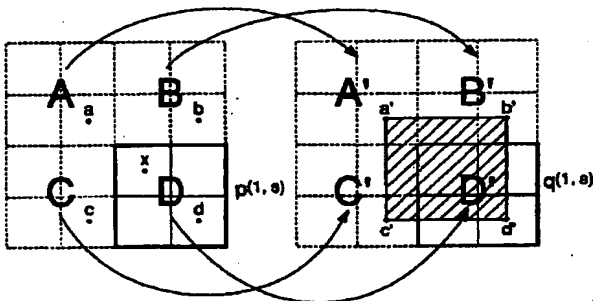


(D)



(C)

【図14】



【図9】

|    |   |  |  |
|----|---|--|--|
|    |   |  |  |
| 10 | 8 |  |  |
| 6  | 3 |  |  |

$p(m, e)$

|   |  |  |
|---|--|--|
| 3 |  |  |
|   |  |  |

$p(m-1, 0)$

|   |  |  |
|---|--|--|
| 8 |  |  |
|   |  |  |

$p(m-1, 1)$

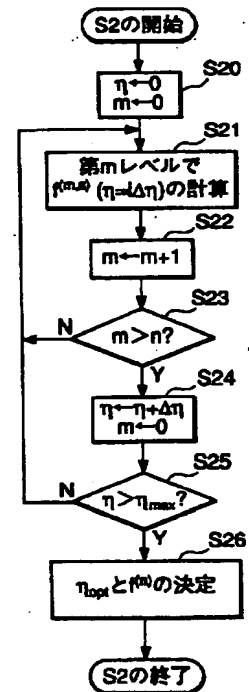
|   |  |  |
|---|--|--|
| 6 |  |  |
|   |  |  |

$p(m-1, 2)$

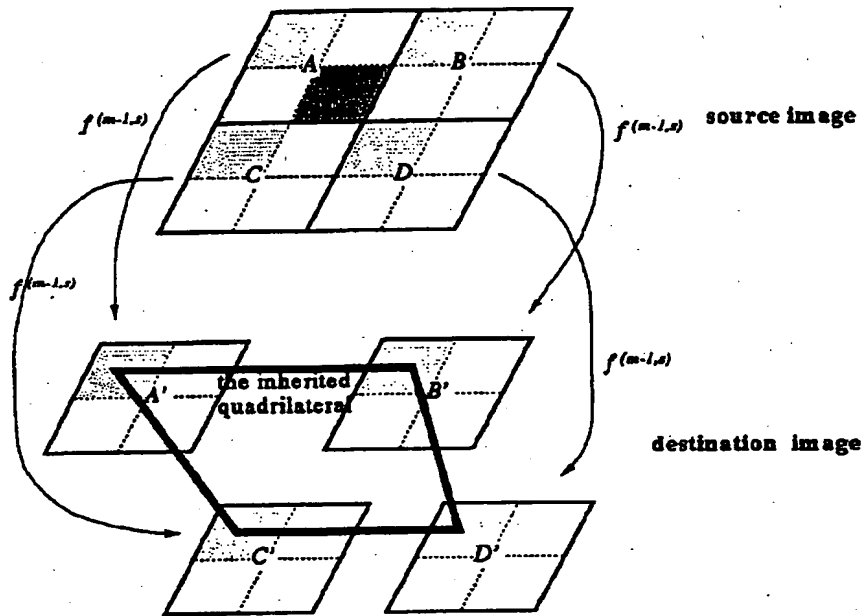
|    |  |  |
|----|--|--|
| 10 |  |  |
|    |  |  |

$p(m-1, 3)$

【図12】

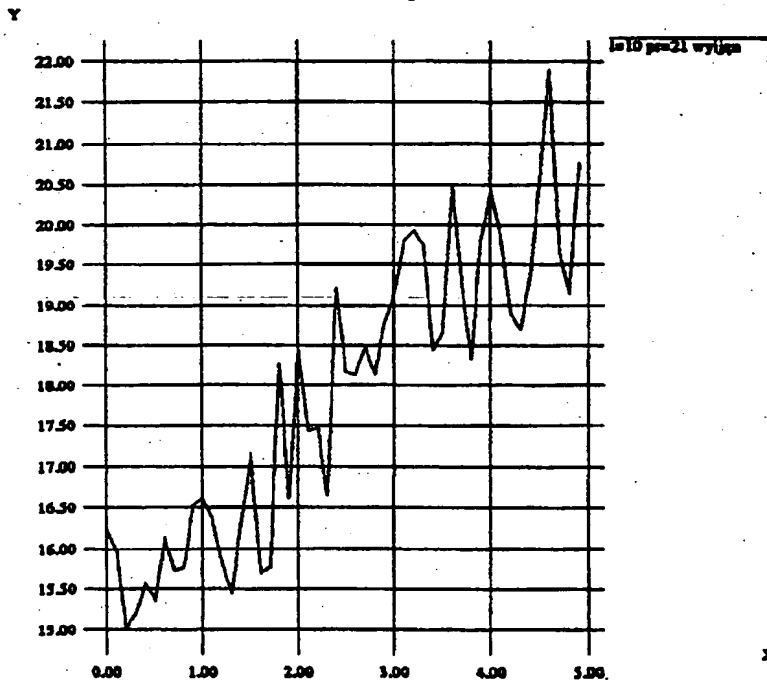


【図3】

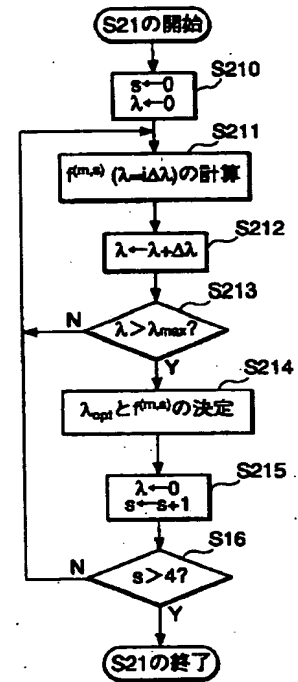


【図4】

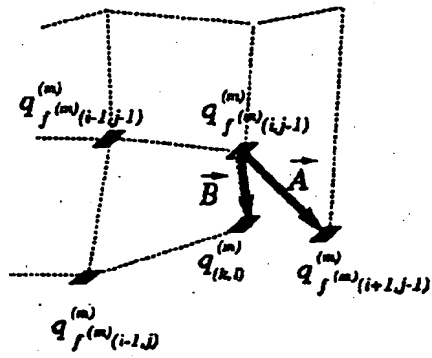
X Graph



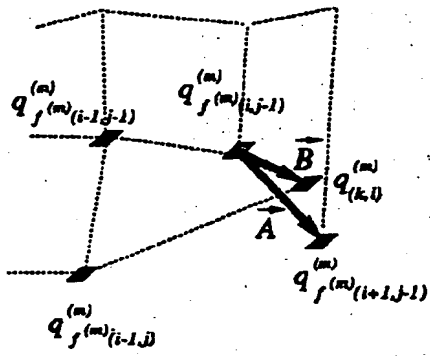
【図15】



【図5】

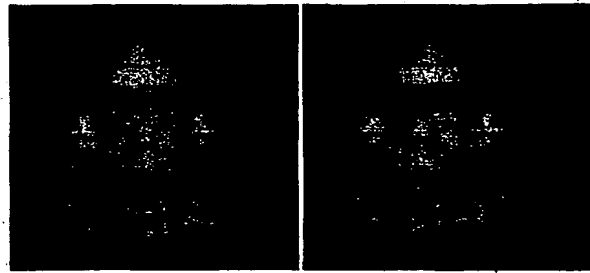


(a.)



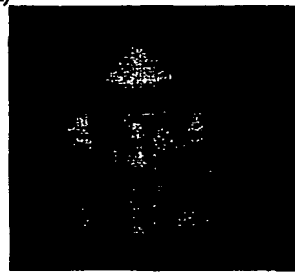
(b)

【図18】



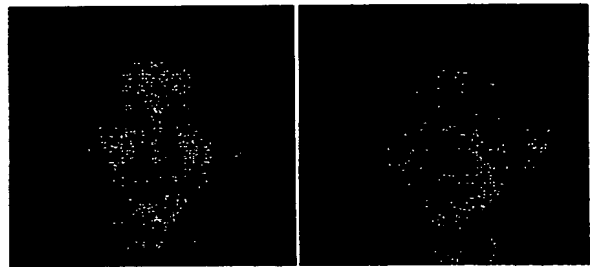
(a)

(b)



(c)

【図19】



(a)

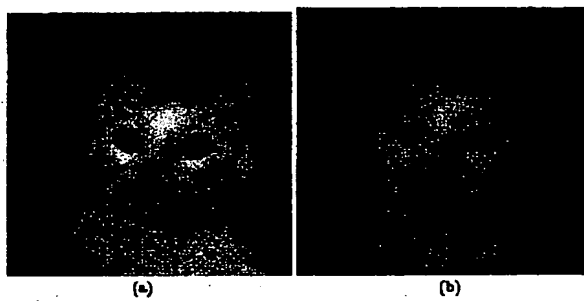
(b)



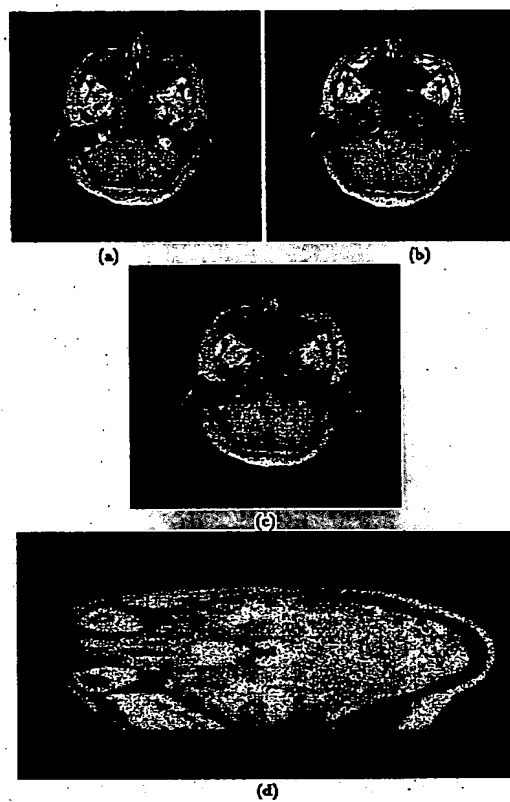
(c)

(d)

【図20】



【図22】





【図21】

カラー写真



(a) (b)



(c)